Journal of Hydrology 556 (2018) 100–118

Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol

Research papers

An improved bias correction method of daily rainfall data using a sliding window technique for climate change impact assessment

P.S. Smitha [a,*], B. Narasimhan [a], K.P. Sudheer [a], H. Annamalai [b]

[a] Department of Civil Engineering, Indian Institute of Technology Madras, Chennai 600036, India
[b] International Pacific Research Center and Department of Oceanography, University of Hawaii, USA

ARTICLE INFO

Article history:
Received 21 August 2017
Received in revised form 29 October 2017
Accepted 7 November 2017
Available online 9 November 2017
This manuscript was handled with the assistance of Ashish sharma, Associate Editor

Keywords:
Hydrological modelling
RCM
Bias correction
Distribution mapping
SWAT

ABSTRACT

Regional climate models (RCMs) are used to downscale the coarse resolution General Circulation Model (GCM) outputs to a finer resolution for hydrological impact studies. However, RCM outputs often deviate from the observed climatological data, and therefore need bias correction before they are used for hydrological simulations. While there are a number of methods for bias correction, most of them use monthly statistics to derive correction factors, which may cause errors in the rainfall magnitude when applied on a daily scale. This study proposes a sliding window based daily correction factor derivations that help build reliable daily rainfall data from climate models. The procedure is applied to five existing bias correction methods, and is tested on six watersheds in different climatic zones of India for assessing the effectiveness of the corrected rainfall and the consequent hydrological simulations. The bias correction was performed on rainfall data downscaled using Conformal Cubic Atmospheric Model (CCAM) to $0.5^\circ \times 0.5^\circ$ from two different CMIP5 models (CNRM-CM5.0, GFDL-CM3.0). The India Meteorological Department (IMD) gridded ($0.25^\circ \times 0.25^\circ$) observed rainfall data was considered to test the effectiveness of the proposed bias correction method. The quantile–quantile (Q–Q) plots and Nash Sutcliffe efficiency (NSE) were employed for evaluation of different methods of bias correction. The analysis suggested that the proposed method effectively corrects the daily bias in rainfall as compared to using monthly factors. The methods such as local intensity scaling, modified power transformation and distribution mapping, which adjusted the wet day frequencies, performed superior compared to the other methods, which did not consider adjustment of wet day frequencies. The distribution mapping method with daily correction factors was able to replicate the daily rainfall pattern of observed data with NSE value above 0.81 over most parts of India. Hydrological simulations forced using the bias corrected rainfall (distribution mapping and modified power transformation methods that used the proposed daily correction factors) was similar to those simulated by the IMD rainfall. The results demonstrate that the methods and the time scales used for bias correction of RCM rainfall data have a larger impact on the accuracy of the daily rainfall and consequently the simulated streamflow. The analysis suggests that the distribution mapping with daily correction factors can be preferred for adjusting RCM rainfall data irrespective of seasons or climate zones for realistic simulation of streamflow.

1. Introduction

Global warming is likely to exert considerable impact on the hydrological cycle (Everett et al., 1996) that subsequently influence water resources. General Circulation Models (GCMs) are widely used to assess the impacts of global warming on time-mean changes in climate and subsequently on the hydrologic cycle. It is well recognized that these GCMs realistically capture large scale circulation patterns and demonstrate high skill in simulating variables which have high degree of spatial uniformity such as surface pressure, wind speed or temperature. The horizontal resolutions at which GCMs operate (~100–200 km) has obvious limitations in resolving and representing small scale features that impact rainfall characteristics and hence moisture related variables from these coarse resolution GCMs cannot be used directly for hydrological/ water resources impact modelling studies in meso/local scales (Hansen et al., 2006; Sharma et al., 2007; Chen et al., 2011; Salvi et al., 2013).

Downscaling techniques have been developed to derive regional information about climate variables from coarse resolution GCM outputs (Themeßl et al., 2011). These techniques take into

* Corresponding author.
E-mail address: pssmithaaa@gmail.com (P.S. Smitha).

https://doi.org/10.1016/j.jhydrol.2017.11.010
0022-1694/

account regional forcings such as orography, lakes and land surface characteristics which influence the local climate (Murphy,1999). There are two approaches to downscale GCM simulations viz. statistical and dynamical downscaling. The statistical downscaling (SD) approach establishes empirical relationship between large scale climate patterns and local climate variables, and incorporate these relationships to GCM outputs to derive regional information (Fowler et al., 2007). In dynamical downscaling (DD) approach, a high resolution Regional Climate Model (RCM) is nested with GCM outputs to derive physically consistent regional weather and climate parameters with sufficient details.

SD and DD methods, both have their strengths and shortcomings. SD methods are more easy to implement and computationally efficient (Hewitson and Crane, 1996). However, it is based on the assumption that statistical relationship between large and local-scale processes remains stationary in future climate (time period). This assumption is perhaps not true given the non-linear and non-stationary nature of the evolving climate. Furthermore, to apply SD needs sufficiently long-term observed data to derive robust statistical relationships for regions of interest. There are several approaches that have been developed over the years to overcome these limitations of SD methods (Hessami et al., 2008; Jeong et al., 2012; Khalili et al., 2013; Ben Alaya et al., 2015; Bhowmik et al., 2017).

DD method on the other hand is complex and difficult to implement. DD uses high resolution RCM model forced with GCM simulated state variables at its lateral boundary along with surface boundary conditions and hence provides finer resolution climate variables (Christensen et al., 2001; King et al., 2012). DD is considered to be very effective in physically characterizing the spatial distribution of climate trends, through process based modelling of the atmospheric system, thus improving simulation of non-stationarity in the climate dynamics (Dell' Aquila et al., 2012) and could improve the inter-annual variability simulated by the GCM by realistically capturing the role of local feedbacks (Guyennon et al., 2013). Hence, DD method is more useful for hydrological model impact studies. However, RCMs are subjected to bias introduced by the systematic errors in the driving GCM variables, along with its own biases due to model formulation (PaiMazumder and Done, 2015). In order to obtain reliable climate simulations, the biases are adjusted using different techniques (Christensen et al., 2008), typically on a monthly scale, which can be used further for hydrological impact studies.

As the study of current climate-hydrology relationships forms the foundation step for future climate change simulations and its effect on water resource management, more reliable and accurate estimates of climate variables become vital. This study aims at providing realistic estimates of rainfall simulations from climate models at daily time scale, by proposing a new methodology for bias correction of RCM derived rainfall data, which is useful for hydrological modelling studies in the context of climate change.

2. Review of literature

The reliability of the RCM output and its further use for hydrologic impact studies are highly dependent on the chosen region of study (Chen et al., 2013). In a chosen region, simulation of RCM rainfall is influenced by regional features such as topography, land-cover/land-use (Salvi et al., 2013), and non-linear interactions between large-scale dynamical and local thermodynamic processes. Rainfall being a highly complex variable, modelling its behaviour at daily time scales, and further at meso/local scales (10–50 km) pose yet another challenge to climate modellers. This challenge is amplified further due to lack of direct observations to constrain and improve model physics. The realistic representation of daily rainfall from climate models play a decisive role in the impact assessment studies, and to assure meaningful runoff simulations from hydrological models, bias correction of climate model outputs is essential (Wood et al., 2004; Eden et al., 2012; Salvi et al., 2013; Chen et al., 2013). Further, in most of the catchments, main factor contributing towards runoff is rainfall. Hence, the selection of a particular bias correction method is more pronounced in the case of rainfall (Haerter et al.,2011; Chen et al., 2013).

In this regard, several bias correction methods have been developed with varying level of complexities to correct the RCM simulations. These include correcting the mean to more advanced methods such as correction of both mean and variance with further advancement in adjusting the quantile values (Sunyer et al., 2014). Linear scaling and local intensity scaling are mean based approaches in which the former correct the monthly means while the latter correct the wet day frequencies and intensities along with the monthly mean (Chen et al., 2013). Power transformation method adjusts the mean and variance of rainfall data (Leander et al., 2008) but not the wet day frequencies and intensities. Distribution mapping or quantile mapping is a distribution based approach which corrects the mean and variance together with wet day frequencies and intensities (Teutschbein and Seibert, 2012).

Li et al. (2010) developed a new method for bias correction called equidistance quantile matching method, based on mixed gamma distribution that accounts for the distribution changes between the baseline and future time periods. Though the results indicated that this method was skilful to incorporate the changed variability and was superior to quantile based mapping method but was applied for monthly rainfall from GCMs. Piani et al. (2010) validated a statistical bias correction method based on distribution mapping method (with Gamma distribution) and the performance was good for seasonal means, heavy rainfall events and seasonal drought index but not for the daily rainfall events. Lafon et al. (2012) compared the performance of linear, non linear, gamma based quantile mapping and empirical distribution-based quantile mapping methods and found out that mean and standard deviation of daily rainfall can be effectively corrected while the correction of skewness and kurtosis of daily rainfall are sensitive to the choice of bias correction method and calibration period. Although, gamma-based quantile mapping method provides better results where the variability in rainfall was captured by gamma distribution, the study employed monthly gamma parameters to correct the daily rainfall data. The performance of distribution derived, parametric and nonparametric transformations were compared by Gudmundsson et al. (2012) and identified that non-parametric transformations possess good proficiency in the reduction of biases in rainfall simulated by RCMs. While assessing hydrological response to climate change, Teutschbein and Seibert (2012) reported that all bias correction methods improved RCM outputs (rainfall and temperature) and distribution mapping method was found to be superior for hydrological simulation but the corrections employed monthly factors. A histogram equalization bias correction method was applied by Argüeso et al. (2013) that could considerably reduce the seasonal and daily biases of rainfall even in areas where the rainfall is over estimated and in areas where the wet-day conditions are not met. As the driving mechanism of rainfall varies throughout the year, the bias correction was applied on a seasonal basis and the gamma parameters were estimated for each season separately. Fang et al. (2015) reported that power transformation and quantile mapping method performed better in correcting standard deviation and percentile values but employed monthly correction factors at daily time scale. To correct the temporal scale biases, Maurer et al. (2010) applied bias correction on a daily basis using the daily average

precipitation (maximum and minimum temperatures) for a particular month while some studies have employed quantile mapping method at daily timescale using a 31 day moving window approach (Themeßl et al., 2011; Thrasher et al., 2012; Wilcke et al., 2013). Further advancements in the area of bias correction with multiple scales and nested approaches at daily scale has been attempted (Mehrotra and Sharma, 2012; Mehrotra and Sharma, 2015).

Another advancement was seen in the deployment of multiple timescales with a three tier cascade bias correction method proposed by Haerter et al. (2011) which can operate on different timescales separately which could reproduce the statistics of the observed data at various timescales. Johnson and Sharma (2012) developed a nested bias correction approach based on the assumption of a linear autoregressive model that could correct the mean, standard deviation and lag-1 autocorrelations of GCM simulations at the monthly and annual time scales which performed well when the biases are not too large. Mehrotra and Sharma (2012) proposed a Recursive Nested Bias Correction approach which is an extension of the nested bias correction which includes five specified time scales and more than three iterations to improve the bias corrected results and was found to be more effective in correcting distribution and persistence attributes at various time scales. Mehrotra and Sharma (2015) developed Multivariate Recursive Nesting Bias Correction approach which simultaneously corrects multiple climate variables at different levels of temporal aggregation to replicate distributional and persistence properties of observed data at multiple time scales to achieve good results across all variables and time scales considered. A further modification of this method known as multivariate recursive quantile nesting bias correction was proposed by Mehrotra and Sharma (2016) which focused at the correction of lag-1 dependence and cross-dependence attributes across multiple time scales and found that the representation of variability and persistence related attributes good. Nguyen et al. (2016) presented a frequency based bias correction method which employs Fourier transform and is independent of the choice of timescale for GCM based monthly precipitation and found that the performance was good for first and second moment statistics and persistence based attributes at multiple time scales.

Recently, much thrust has been given to bivariate distribution of rainfall and temperature for correcting their joint distribution which could reduce the biases not only in the mean and variance, but also in the correlation between the two variables (Li et al., 2014). A multivariate empirical copula–bias correction (EC–BC) proposed by Vrac and Friederichs (2015) combines one dimensional bias correction and reshuffling of multivariate spatiotemporal data based on the rank structure derived from training data to obtain realistic inter-variable, spatial and temporal dependencies but makes stationarity assumption which could only replicate patterns derived from historical series. A bias correction method which combines principal components and quantile mapping was attempted by Hnilica et al. (2017) which takes into account correction of multivariate data sets and could substantially reduce the bias in the distribution of individual variables, covariance and correlation structures of daily precipitation. But spatial dependence structure of daily precipitation was only corrected and the dependence between climate variables was not attempted. Cannon (2016) developed a multivariate bias correction (MBC) algorithm which is a multidimensional analog of univariate quantile mapping method and corrects the dependence structure of Pearson correlation and Spearman rank correlation to match the observed distribution. The performance of MBC was much convincing compared to the univariate quantile mapping method but was applied on monthly time series. A further improvisation of multivariate bias correction method was adapted (Cannon, 2017) by implementing N-dimensional probability density function transform algorithm (MBCn) which could transfer all characteristics of observed multivariate distribution to the corresponding multivariate distribution of climate projections. Though the results were closely matching with the observed data for spatio-temporal autocorrelation of rainfall data, the method possess high computational complexities. However, Dosio and Paruolo (2011) reported that univariate bias correction methods have prominence over multivariate methods when a good measure of dependence among climatic variables are accounted for. Moreover, rarely multivariate approach aims at conserving the dependence between more than two variable (Hempel et al., 2013).

Even though immense progress has been achieved in recent years in the development and assessment of bias correction methods, temporal variation of rainfall was accounted only in few studies while correcting the daily RCM rainfall. Many of the existing bias correction methods which are performed on a monthly basis by grouping data into months or seasons often result in masking the characteristics of daily rainfall (Kim et al., 2015). The day to day variability in the rainfall characteristics are mainly caused by the magnitude of diurnal cycle and local weather/synoptic system (Haerter et al., 2011). Hence, when a bias correction is performed based on correction factors estimated month/seasonal basis and applied on a daily timescale, the variations of daily rainfall cannot be accurately incorporated due to sharp transition in the bias correction factors between one month/season and another month/season. This sharp transition in bias correction factor can distort the daily characteristics of rainfall. Moreover, for climate change impact studies, most of the hydrological models require good quality rainfall data at daily time scale for obtaining realistic flow simulations. Hence, an improved bias correction method which accounts for variability of rainfall at daily time scale is vital for use in climate change studies and has been focused in the present study. Therefore, in this study we propose a method to account for the temporal variability in daily rainfall, by performing bias correction of daily rainfall with daily correction factors using a sliding window technique. The effectiveness of the proposed method is evaluated for different climatic zones across India and is shown to perform better than those methods that depend on monthly correction factors. Further, the impact of proposed method on hydrologic simulation of river discharges using Soil and Water Assessment Tool (SWAT) (Arnold et al.,1998) is evaluated for six catchments in India, each located in different climatic zones.

3. Study area and climate datasets

The study area (Fig. 1) consists of six catchments; *Achankovil*, *Vaippar*, *Birhidhang*, *Ken*, *Subarnarekha* and *Sabarmathi* river basins of India, each representing different regional climate and the details of which are given in Table 1. Specifically, these regions were chosen based on the timing, amount and monsoonal characteristics (South West or North East) of rainfall.

i) *Achankovil River* located in the state of Kerala originates from the Western Ghat mountain range in the southern peninsula. The basin receives rainfall during both South West/summer monsoon (June to September) as well as North East/retreating monsoon (October to December).
ii) *Vaippar River* also originates from the Western Ghat mountain range, but on the leeward side (rain shadow region), located in the state of Tamil Nadu. The rainfall in the basin is characterized by North East monsoon accounting for the majority of the rainfall.
iii) *Birhidhang River*, which is an important south bank tributary of Brahmaputra river flows through the states of Assam and Arunachal Pradesh. The basin receives heavy monsoonal rain

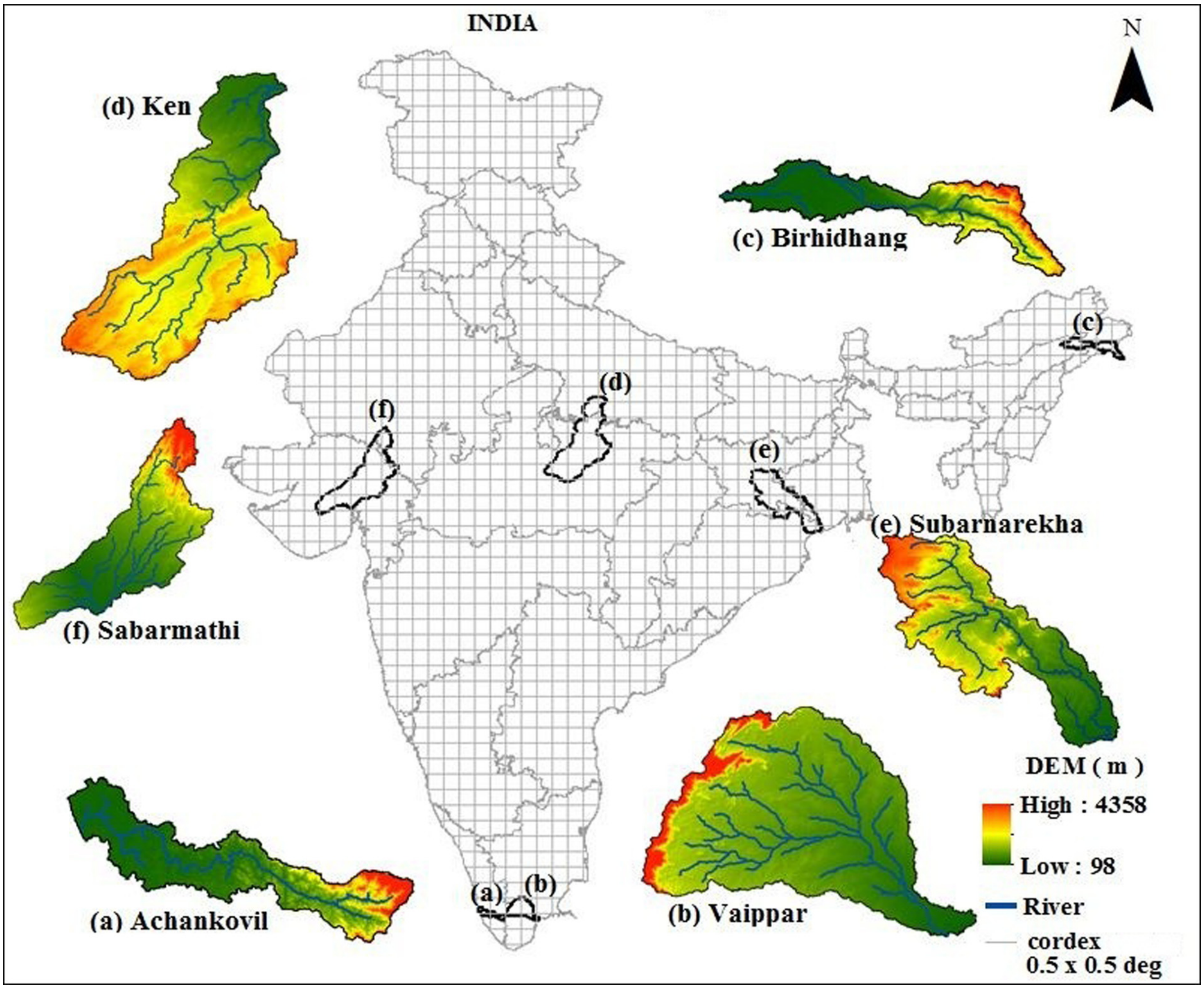

Fig. 1. Location map of the study area.

Table 1
Characteristics of the study areas.

Catchments	Area (km^2)	Length (km)	Latitudinal extent	Longitudinal extent	Mean annual rainfall (mm/year)	Climate Zone (Koppen)
Achankovil	1484	124	9°2′–9°19′N	76°24′–77°18′E	2600	Tropical wet
Vaippar	5431	128	8°57′–9°47′N	77°16′–78°21′E	770	Semi arid
Birhidhang	6000	332	27°5′–27°47′N	95°6′–97°10′E	3000	Humid Subtropical (no dry season)
Ken	28058	427	23°12′–25°54′N	78°30′ –80°36′E	1150	Humid Subtropical
Subarnarekha	19296	395	21°33′–23°32′N	85°9′–87°27′E	1800	Tropical wet and dry
Sabarmathi	21674	371	22°15′–24°53′N	72°15′–73°49′E	750	Arid

during South West monsoon. However, the basin also receives intermittent thundershowers during the dry seasons of March to May. The basin experiences cold, humid winter with shorter summer.

iv) *Ken River* originates from north west slopes of Kaimur hills, flows through the states of Uttar Pradesh and Madhya Pradesh before joining river Yamuna. The basin which is fed mostly by the South West monsoon (90% of annual rainfall) is situated in subtropical humid climatic zone of India and experiences dry winter and very dry summer.

v) *Subarnarekha River* originates from Chotanagpur plateau flows through the states of Jharkhand, West Bengal and Odissa before its confluence with Bay of Bengal. The basin is mostly influenced by the South-West monsoon and receives 90% of rainfall during June-October. The basin experiences a humid sub tropical climate with hot summer and mild winters.

vi) *Sabarmathi River* originates from Aravalli hills of Rajasthan. The basin is located in the hot arid regions of western India and experiences hot summer with temperature ranges of 42 °C–45 °C and moderate winter. The basin receives its 95% of the annual rainfall during the monsoon months of June-September. The river flows through states of Rajasthan and Gujarat before draining into Gulf of Cambay of Arabian Sea.

This study uses a high resolution (0.25° x 0.25°) gridded daily rainfall data obtained from India Meteorological Department (IMD) for the period 1901–2013 (Pai et al., 2014). As temperature data at the same resolution was not available, gridded daily temperature data at the resolution of $1^\circ \times 1^\circ$ (latitude/longitude) developed by IMD for the period 1969–2005 was used (Srivastava et al., 2009). The two best performing GCMs for the Indian subcontinent were selected for this study based on the model evaluation works by Sperber et al. (2012), Sengupta and

Rajeevan (2013), and Hasson et al. (2014). Sperber et al. (2012) assessed the performance of CMIP3 and CMIP5 models in simulating the Asian summer monsoon. Based on statistical metrics, the study concluded that GCMs fidelity in simulating the Indian monsoon is still low. However, they identified that CNRM-CM5.0 and GFDL-CM3.0 models realistically simulate the June-September summer rainfall climatology over the Indian monsoon region. The current study, thus adopts CNRM-CM5.0 and GFDL-CM3.0 models of CMIP5 versions. The data from these two GCMs downscaled to $0.5^\circ \times 0.5^\circ$ using the RCM Conformal-Cubic Atmospheric Model (CCAM) for the reference period (1970–2005) was obtained from Coordinated Regional Downscaling Experiment (CORDEX) for South Asia from the Centre for Climate Change Research (CCCR), Indian Institute of Tropical Meteorology (http://cccr.tropmet.res.in/cordex/). For this study, the downscaled RCM rainfall data was used for the evaluation of various bias correction methods.

The current study used 90 m digital elevation model (DEM) from the Shuttle Radar Topography Mission (SRTM) for delineating the watershed. The land cover data developed by National Remote Sensing Centre (NRSC), India from AWiFS sensor at 56 m resolution was used. The soil map used for hydrologic modelling was Harmonized World Soil Database v 1.2 (Nachtergaele et al., 2008), obtained from the Food and Agriculture Organization (FAO) of the United Nations.

4. Methodology

The existing approaches used for bias correction of rainfall data aims at correcting bias in RCM simulations by employing a transformation algorithm to correct daily values based on statistics and data pooled for each month. Instead, in this study we are proposing to apply these transformation algorithms for correcting daily rainfall values based on statistics from data pooled for a 31 day sliding window. The basic principle involved in the estimation of daily correction factors is to employ a sliding window approach for each of the 365 days in a year over the available number of years of observed rainfall data in order to choose a minimum of 100 rainy day values so as to ensure a good number of data points for obtaining probability distribution plots. For this, window sizes ranging from 10 to 40 days around each of the 365 days in a year were tested and it was observed that a minimum window size of 31 days (centred around each day of the year) was acceptable in obtaining a set of at least 100 rainy days from the 30 year record during dry months across different study areas. These selected rainy day values were further used in the various bias correction methods to estimate the daily correction factors which were used for corrections of RCM simulations. For this, rainfall data from RCM grids were bias corrected with respect to that of the nearest IMD gridded rainfall data. The bias correction methods based on both daily and monthly correction factors are evaluated based on Nash Sutcliffe model efficiency Coefficient (NSE) and quantile- quantile (Q-Q) plots with respect to IMD data.

Prior to bias correction, the extent to which various RCMs capture the regional distribution of the monsoon and rainfall pattern was analysed based on the frequency and magnitude of rainfall. The data obtained from IMD was used as observed data to assess the fidelity of the RCMs in reproducing rainfall characteristics of current climate. As the comparison of number of rainy days and rainfall volume provides a simple and effective way to assess the performance of CMIP5 models with observed data, it was used to assess the effectiveness of the selected models in representing the climate of the study area. The effect of bias correction methods on hydrological model simulations are studied based on long-term median hydrographs, flow duration curves and streamflow quantiles for six catchments with different climatology using two CMIP5 models (GFDL and CNRM). The estimation of different correction factors for different bias correction methods and their performance were all evaluated for the same time period 1974–2005.

4.1. Bias correction methods

In this study, five bias correction methods based on daily correction factors have been tested, including a modification proposed to the power transformation method. The details of the methods are explained in the following sections:

4.1.1. Linear scaling (LS) of rainfall

Lenderink et al. (2007) proposed linear scaling approach of bias correction based on monthly correction values. The rainfall is corrected with a multiplicative ratio of observed monthly mean to the uncorrected data (RCM data for the baseline/historic run). In this study the correction is based on multiplier term on a daily basis using a 31-day sliding window.

$$P_{cor,d} = P_{RCM,d} \times \frac{\mu(P_{obs,d})}{\mu(P_{RCM,d})} \tag{1}$$

where $P_{cor,d}$ and $P_{RCM,d}$ are corrected and uncorrected rainfall respectively on day 'd' and $\mu(\cdot)$ represents the expected value for the 31-day sliding window for the day 'd'.

4.1.2. Local intensity scaling (LOCI) of rainfall

Many a time the climate models (GCM/RCM) rain frequently (drizzle effect) albeit in small quantities, which may skew the statistics used for bias correction. Hence, Schmidli et al. (2006) presented local intensity scaling which could adjust the mean as well as both wet day frequencies and wet day intensities of rainfall time series. Threshold rainfall ($P_{th,d}$) is fixed such that the number of rainy days above 2.4 mm (definition of a light rainy day as per India Meteorological Department) in the observed and the simulated records are of similar length over a 30-year period. Further, RCM rainfall is corrected by redefining days with rainfall less than $P_{th,d}$ as dry days with no rainfall, and for rainfall more than $P_{th,d}$ the threshold value is subtracted from the RCM rainfall. The linear scaling factor s_d is estimated by the following ratio:

$$s_d = \frac{\mu(P_{obs,d}|P_{obs,d} > 0)}{\mu(P_{RCM,d}|P_{RCM,d} > P_{th,d}) - P_{th,d}} \tag{2}$$

The correction factor thus derived is applied to the rainfall data as given in Eq. (3).

$$P_{cor,d} = \begin{cases} 0, & \text{if } P_{RCM,d} < P_{th,d} \\ (P_{RCM,d} - P_{th,d}) \times S_d & \text{otherwise} \end{cases} \tag{3}$$

In this study, a slight modification of existing LOCI method was applied which includes rainfall thresholding prior to multiplying linear scaling factor s_d, and $P_{th,d}$ was calculated on a daily basis for a 31 day sliding window. This could remove most of the very light rainy days.

4.1.3. Power transformation (PWTR) of rainfall

The power transformation method uses an exponential form (aP^b) to adjust the variance of rainfall series (Leander and Buishand, 2007; Leander et al., 2008). The parameter 'b' is calculated on a daily basis for a 31-day sliding window by matching the co-efficient of variation of observed and RCM daily rainfall series. The exponential term 'b' for each day is estimated, which minimizes the function:

$$f(b_d) = \frac{\sigma(P_{obs,d})}{\mu(P_{obs,d})} - \frac{\sigma\left(P_{RCM,d}^{b_d}\right)}{\mu\left(P_{RCM,d}^{b_d}\right)} \tag{4}$$

where b_d is the exponent for the dth day, σ represents the standard deviation and μ represents the mean of the rainfall. $P_{obs,d}$ and $P_{RCM,d}$ indicates the observed rainfall and uncorrected RCM rainfall for the baseline scenario for the day 'd'. Following which, an intermediary corrected rainfall, $P^*_{RCM,d}$ is estimated. Thereafter, the scaling parameter, s_d which is the ratio of observed mean and mean of the intermediary corrected rainfall is calculated. The final corrected rainfall series, P_{cor} is obtained by multiplying the intermediary corrected rainfall with the scaling parameter.

$$P^*_{RCM,d} = P^{b_d}_{RCM,d} \tag{5}$$

$$s_d = \frac{\mu(P_{obs,d})}{\mu\left(P^*_{RCM,d}\right)} \tag{6}$$

$$P_{cor} = P^*_{RCM,d} \times s_d \tag{7}$$

4.1.4. Modified power transformation (MPWTR) of rainfall

The Power transformation method does not correct the biases in the wet day frequencies and intensities (Teutschbein and Seibert, 2012), caused by the "drizzle effect" of the models. Hence, we tested a hybrid method of combining rainfall thresholding (similar to LOCI) prior to applying correction by power transformation method. At first, rainfall threshold was calculated for the RCM rainfall data as explained in LOCI method (Section 4.1.2) so that the wet day frequencies are similar in RCM as those in IMD. Thereafter, RCM rainfall was corrected by redefining days with rainfall less than $P_{th,d}$ as dry days with no rainfall and for rainfall more than $P_{th,d}$ the RCM rainfall was shifted by $P_{th,d}$. The equations employed for modified power transformation are given as in Eqs. (8)–(11).

$$f(b_d) = \frac{\sigma(P_{obs,d})}{\mu(P_{obs,d})} - \frac{\sigma(P_{RCM,d} - P_{th,d})^{b_d}}{\mu(P_{RCM,d} - P_{th,d})^{b_d}} \tag{8}$$

$$P^*_{RCM,d} = (P_{RCM,d} - P_{th,d})^{b_d} \tag{9}$$

$$s_d = \frac{\mu(P_{obs,d})}{\mu(P^*_{RCM,d})} \tag{10}$$

$$P_{cor} = P^*_{RCM,d} \times s_d \tag{11}$$

4.1.5. Distribution mapping (DM) of rainfall

The distribution mapping method corrects the distribution function of RCM simulated rainfall in such a manner that it agrees with the distribution function of observed time series (Quantile-Quantile mapping or Q-Q mapping). The shifting of distributions to correct the RCM simulations with respect to observed rainfall can be achieved with the help of a transfer function (Sennikovs and Bethers, 2009). For each site, various probability distributions such as Gamma, Gumbel, Weibul, Normal and Lognormal were applied to the rainfall series and were tested for the goodness of fit using Kolmogorov-Smirnov (K-S) test. In the present study, Gamma distribution was found to be the best fitting distribution for rainfall data, from different climate and for different seasons throughout the year. Gamma distribution function was constructed for RCM simulated rainfall values with shape parameter α and scale parameter β as given in Eq. (12):

$$f(x|\alpha,\beta) = x^{\alpha-1} \times \frac{1}{\beta \times \Gamma(\alpha)} \times e^{\frac{-x}{\beta}};\ x \geqslant 0;\ \alpha,\beta > 0 \tag{12}$$

As the higher number of rainy days in the simulated record due to the "drizzle effect" could skew the fitted distribution, rainfall thresholding as explained in Section 4.1.2 was subtracted from the RCM rainfall so as to remove the days with very low rainfall. The shape and scale parameters for all the 365 days were found out for the observed and RCM simulated rainfall on a daily basis using a 31-day sliding window centred on day 'd'. The inverse of gamma distribution function was found out for these gamma cumulative distribution function values with the corresponding day's shape and scale parameter of observed values. This value is taken as the corrected RCM rainfall for the current climate condition. This procedure of Q-Q mapping can be expressed mathematically as:

$$P_{cor,d} = F_\gamma^{-1}\left(F_\gamma\left(P_{RCM,d}|\alpha_{RCM,d},\beta_{RCM,d}\right)|\alpha_{obs,d},\beta_{obs,d}\right) \tag{13}$$

4.2. Impact of bias correction on hydrologic model simulation of river discharges

The uncertainties introduced by the climate and hydrological models combined with the natural climate variability (Muerth et al., 2013) affects the amount of simulated streamflow from the catchment. In this study, to illustrate the effect of various bias correction methods on the hydrologic response of a catchment, the hydrological model SWAT was employed. The SWAT model parameters were not calibrated as the intention was to understand the effect of different bias correction methods of rainfall on the hydrologic response of the catchment, while keeping the SWAT model parameters the same between simulations. So the only difference among the simulations of a river basin is the rainfall data from different bias correction methods. The rainfall and temperature data of the selected downscaled GCMs were directly used to run the SWAT model without any bias correction and then with the various bias corrected rainfall data. This was then compared with SWAT simulations with IMD gridded weather data (observed). This helps in assessing whether the bias correction methods applied to downscaled rainfall data, when used for hydrological modelling are able to simulate streamflows comparable to those simulated using IMD gridded rainfall data. SWAT model simulations were performed for the time period from 1970 to 2005, with a model warm up period of four years. Hence, only the streamflow simulations from 1974 to 2005 was used for further analysis. The proposed methodology was applied on six river basins located on different climate zones within India. The effect of bias correction methods on hydrological model simulations were studied based on comparing long-term median monthly hydrographs and flow duration curves simulated using bias corrected rainfall data compared to those obtained using simulations with IMD gridded weather data.

5. Results and discussions

The temporal characteristics of observed and simulated number of rainy days and rainfall magnitude above 2.4 mm (definition of a light rainy day as per India Meteorological Department) during each month (long-term monthly average) before bias correction were compared to assess the performance of selected GCMs in six basins falling in different climatic zones (Fig. 2). The results of average number of rainy days and rainfall amounts for uncorrected RCM data are presented for one representative RCM grid (0.5° × 0.5°) for each basin.

Analysis of the total number of rainy days simulated in a year (over a 30 year simulation period) by both CNRM-CM5.0 and GFDL-CM3.0 compares reasonably well with average number of rainy days from IMD gridded data for all catchments except Vaippar basin which is located in the semi-arid zone/rain shadow region. However, there are considerable differences in the average number of rainy days simulated in all the basins on a monthly basis (Fig. 2), including the month with highest number of rainy days.

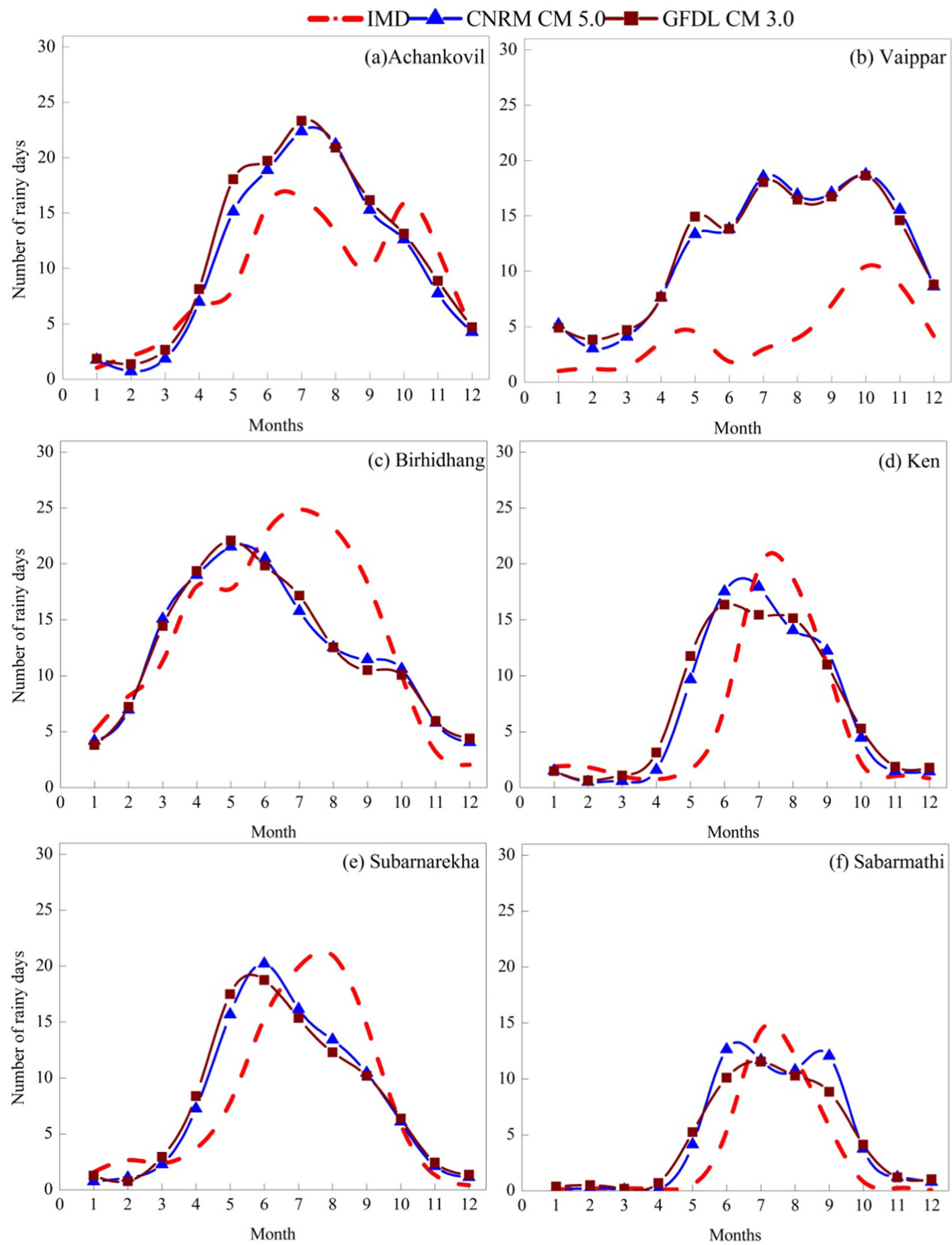

Fig. 2. Uncorrected RCM simulated rainy days by CNRM-CM5.0 and GFDL-CM3.0 in (a) Achankovil, (b) Vaippar, (c) Birhidhang, (d) Ken, (e) Subarnarekha and (f) Sabarmathi basin.

The inter model variability between the two best performing climate models is very minimal (Figs. 2 and 3).

As mentioned earlier, the Achankovil basin, which is on the windward side of the Western Ghats, experiences two rainfall seasons, South-West monsoon and North-East monsoon (Fig. 2(a)). However, the GCMs seem to simulate primarily the South-West monsoon, which is a major rainfall season for much of the Indian sub-continent. The North-East or retreating monsoon caused due to high pressure because of the cooling of landmass over the Himalayas and the Indo-Gangetic plain after the autumnal equinox (September). This brings air mass laden with moisture from Bay of Bengal and northern Indian Ocean resulting in substantial rainfall in several parts of southern peninsula (Kerala and Tamil Nadu). Even the best performing GCMs selected for this region, do not seem to capture the spatial and temporal aspect of this retreating monsoon. This could be observed even in the simulation of rainy days for Vaippar basin (on the leeward side of the Western Ghats, east of Achankovil basin), which dominantly receives rainfall only during the North East monsoon (Fig. 2(b)). However, both the GCMs simulate higher rainy days in Vaippar basin even during the South West monsoon (2–3 times higher than observed number of rainy days; perhaps due to the problem of drizzle effect in the

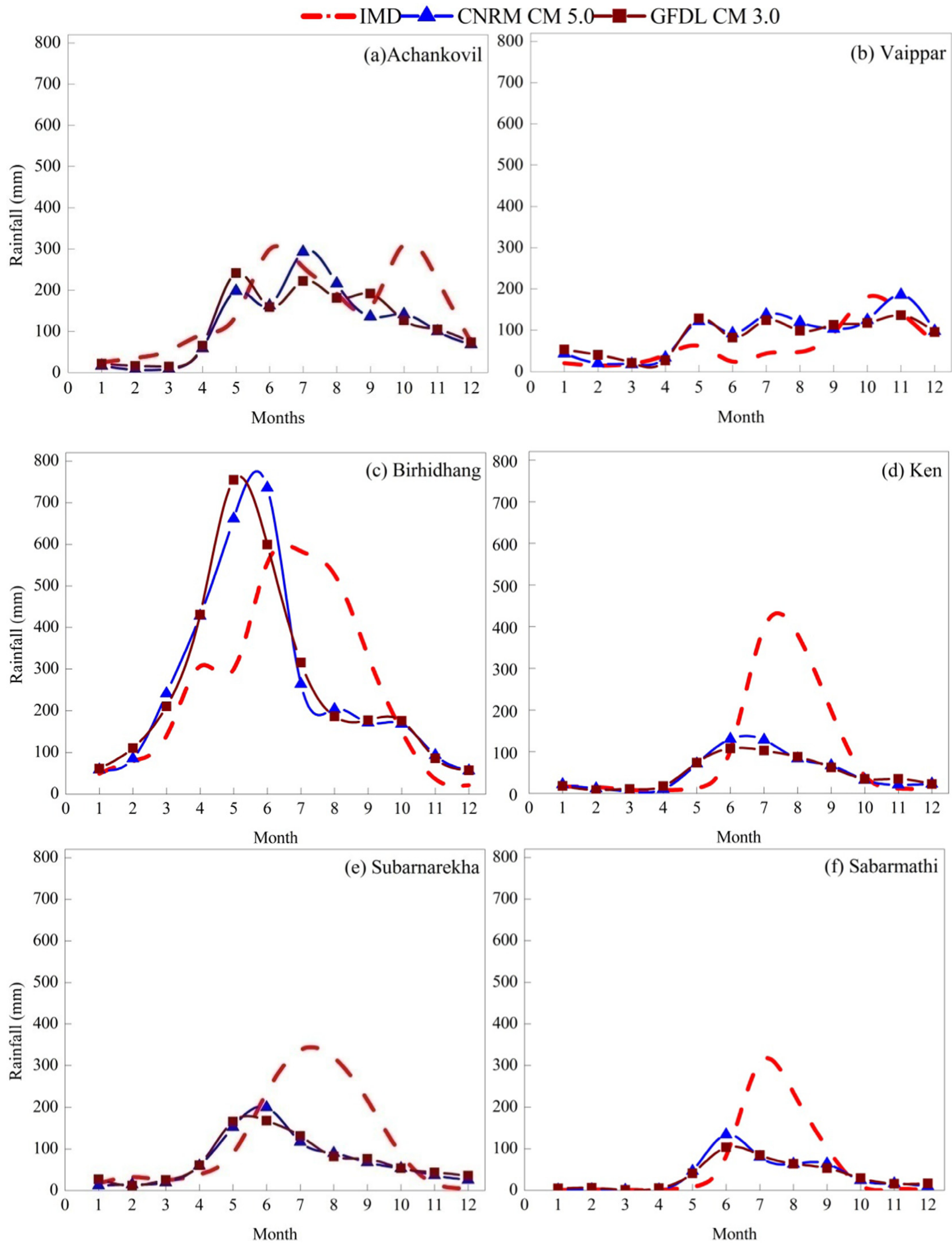

Fig. 3. Uncorrected RCM simulated rain volume by CNRM-CM5.0 and GFDL-CM3.0 in (a) Achankovil, (b) Vaippar, (c) Birhidhang, (d) Ken, (e) Subarnarekha and (f) Sabarmathi basin.

models in this region). In terms of average monthly rainfall magnitude, for Achankovil basin, both the GCMs seem to reasonably capture the magnitude of South West monsoon rainfall, but not the North East monsoon (Fig. 3(a)). For Vaippar basin, even though the number of rainy days in the North East monsoon was over estimated by the GCMs, the total rainfall magnitude was simulated well, with an overestimation of rainfall magnitude during the South West monsoon (Fig. 3(b)).

Birhidhang basin is located in the north-east corner of India in humid-subtropical climate without any dry season. The dominant rainy season is the South West monsoon with peak rainy days and rainfall quantities during the month of July (Figs. 2(c) and 3(c). Compared to observations, though the rainy days and rainfall quantities are reasonably simulated by the climate models, the simulated peak monsoon season is ahead by two months, which clearly indicates the lacuna of even the best climate models in capturing the regional processes.

Ken basin located in central India originating in the Kaimur hills with humid-subtropical climate experiences peak monsoon season during the month of July. The climate models seem to simulate the rainy days reasonably well, but with a one month shift (advance) in the peak monsoonal occurrence (Figs. 2(d) and 3(d)). However, the

rainfall magnitudes were highly underestimated by both the best performing climate models. This again indicates the lacuna of the existing climate models.

Subarnarekha basin is located in eastern India and experiences a tropical wet and dry climate. The dominant rainy season is the South West monsoon with peak rainy days and rainfall quantities during the month of July (Figs. 2(e) and 3(e)). Although the rainy days are reasonably simulated by the climate models, the models simulate the peak monsoon season two months ahead of the actual monsoon peak (similar to Birhidhang basin), which clearly indicates the lacuna of even the best climate models in capturing the Indian summer monsoon. The climate models also underestimate the rainfall magnitude by as much as 50% of the observed rainfall over the 30 year record.

Sabarmathi basin is located in western India and experiences an arid climate with a dominant rainy season of South West monsoon with peak rainy days and rainfall quantities during the month of July (Figs. 2(f) and 3(f)). The rainy days are reasonably simulated by the climate models, however the models clearly underestimate the rainfall quantities by as much as 60% of observed rainfall.

In summary, the climate models tend to capture the peak rainfall during the South West monsoon season earlier by as much as one to two months. Although the number of rainy days simulated in the regions dominated by the South West monsoon is comparable to the observed data, the models seem to underestimate the rainfall magnitudes by as much as 50% in many cases, except in the north east portion (Birhidhang basin) where there is a slight over estimation of rainfall. Further, in the regions experiencing north east monsoon (retreating monsoon) in the southern peninsula, the models exhibit extreme "drizzle effect". This is apparent in the Vaippar basin, where the number of rainy days simulated by the models is as much as two to three times higher than the observations, while the rainfall magnitudes are comparable. At the moment, this lacuna exhibited by the climate models can only be rectified by bias correction techniques, while the science of climate modelling is progressing to improve the model representations.

5.1. Performance evaluation of bias correction methods

5.1.1. Q-Q plots of bias correction methods with daily correction factors

Q-Q plots provide useful comparison of the response of rainfall distribution across various bias corrected rainfall values. In order to compare the overall performance of each bias correction method based on daily correction factors, we first evaluated the quantiles of observed, uncorrected and bias corrected RCM simulations for a chosen day for both the CMIP5 models for selected grids of all six catchments. This was done to evaluate whether the daily distribution of observed rainfall could be replicated by this new bias adjustment methodology based on daily correction factors. As a demonstration of this, Q-Q plot of CMIP5 (CNRM-CM5.0, GFDL-CM3.0) models for Achankovil basin on the 160th day (31 day window centred on 160th day during the peak of South West monsoon) is shown in Fig. 4. As indicated by the Fig. 4(a) and (b) all the bias correction methods based on daily correction factors consistently show improved performance for both the CMIP5 models in Achankovil basin with high NSE values (above 0.9).

Fig. 4(a) and (b) indicates that the performance of distribution mapping (DM), Modified Power Transformation (MPWTR) and Local intensity scaling (LOCI) methods closely followed the IMD values (observed rainfall) for GFDL-CM3.0 in Achankovil basin. PWTR method underestimated the high rainfall values for GFDL-CM3.0 model. While for CNRM-CM5.0 model all the bias correction methods followed the IMD values, with DM method performance better than the others. The same procedure was repeated for other catchments also and it was observed that in all the six basins, daily based bias correction methods greatly improved the rainfall values when compared to IMD values. Hence Q-Q plot for a particular day could clearly bring out the improvement achieved by various bias correction methods based on daily correction factor, even though the performance of different methods varied depending on rainfall parameters adjusted (eg. mean, intensity, variance or distribution of rainfall).

Bias correction of daily rainfall was also carried out using monthly bias correction factors (without the sliding 31-day window) as detailed in Teutschbein and Seibert (2012). The Q-Q plot for the 160th day for the same grid based on monthly correction factors applied on a daily basis (Fig. 5), show a significantly lower performance with NSE values varying from 0.56 to 0.84 for different methods when compared to the Q-Q plots based on daily correction factors. Thus the results (Figs. 4 and 5) indicate that, irrespective of the bias correction method, the daily correction factors could considerably improve the bias correction when compared to the monthly correction factors.

5.1.2. Comparison of performance statistics of bias correction methods with daily and monthly correction factors

Statistical evaluation of bias correction methods with daily and monthly correction factors was carried out for all the catchments using the performance measure Nash Sutcliffe Efficiency (NSE). Based on the Q-Q comparison as explained in the previous section, NSE was estimated for each of the 365 days in a year for the entire dataset (for various bias correction methods) and uncorrected data with reference to the IMD gridded rainfall data (observed). The distribution of NSE values for each grid for 365 days was assessed with Box and Whisker plots.

For brevity and illustration purpose, the box plot comparison of NSE values only for Achankovil basin (Fig. 6) is shown as a representative plot. The NSE values obtained for monthly correction factors based bias correction method was considerably lower than that of the daily based method and also vary over a very broad range. This observation is also consistent across other catchments. It could be seen from the box plot (Fig. 6(a)) that for LOCI, MPWTR and DM method, the NSE values are high for both the CMIP5 models (CNRM-CM5.0 and GFDL-CM3.0). Although, the NSE values of PWTR for GFDL-CM3.0 model are high but are distributed across abroad range of values. As seen from the box plot (Fig. 6(b)) the NSE values are considerably lower across all bias correction methods by adopting the correction factor on a monthly basis as compared to applying correction factor on a daily basis as proposed in this study. This may be due to the fact that the choice of an appropriate time scale of correction factor helps in reducing the discontinuities of daily rainfall caused due to sharp changes in bias correction factors between months. Chen et al. (2013) has also reported that bias correction factors calculated on a monthly scale has a limited effect on properties when used for correction at the daily scale. Another drawback of using monthly time scale of correction factors is that, weather extremes which can occur anytime over a month cannot be corrected effectively (Pierce et al., 2015). Hence, adopting daily bias correction factors seem to provide good improvement in the correction of the rainfall data from climate models prior to their use in hydrologic models.

In addition to this, long-term mean monthly rainfall volume of all bias correction methods for all the six catchments was evaluated with those of IMD gridded rainfall data. The mean absolute error (MAE) was used as an evaluation statistic. From Table 2, it is clear that the MAE of DM method is consistently lower among all the bias correction methods irrespective of bias adjustment based on monthly or daily correction factors. This indicates the superior performance of the DM method for bias correction. Further, from Table 2, it is very clear that irrespective of the bias

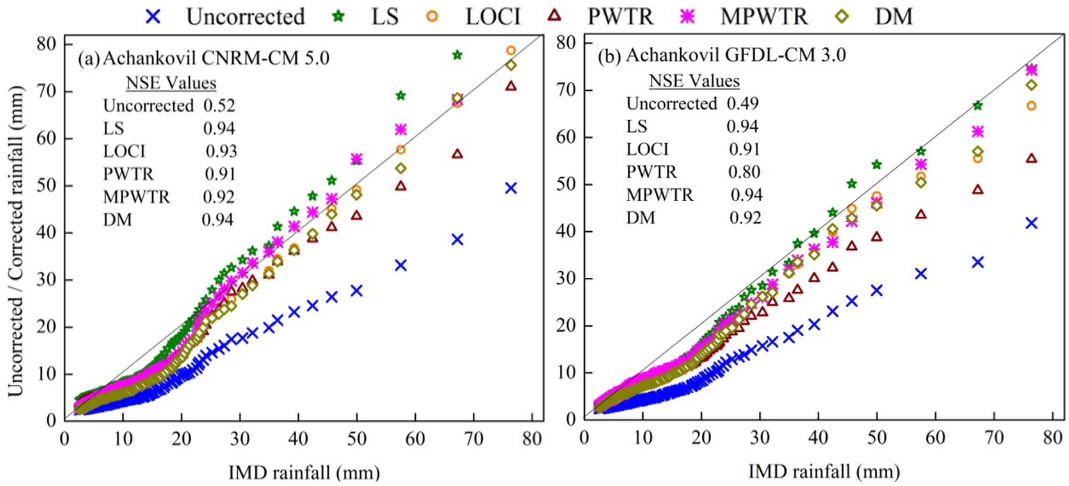

Fig. 4. Quantile-Quantile plots of Achankovil basin based on daily correction factor for (a) CNRM-CM5.0, (b) GFDL-CM3.0 [Uncorrected – Downscaled data; LS – Downscaled data corrected using Linear Scaling; LOCI – Local Intensity Scaling; PWTR – Power Transformation; MPWTR – Modified Power Transformation; DM – Distribution Mapping].

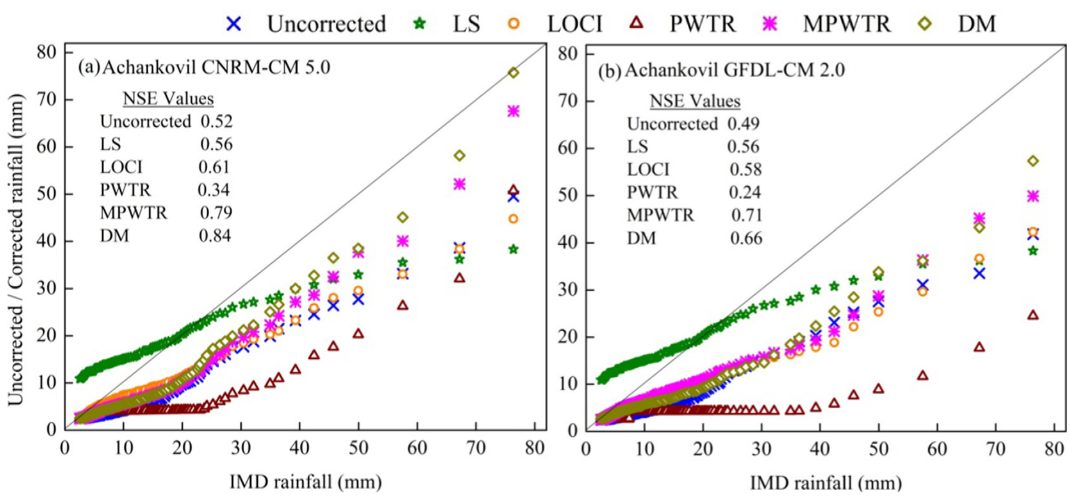

Fig. 5. Quantile-Quantile plots of Achankovil basin based on monthly correction factor for (a) CNRM-CM5.0, (b) GFDL-CM3.0 [Uncorrected – Downscaled data; LS – Downscaled data corrected using Linear Scaling; LOCI – Local Intensity Scaling; PWTR – Power Transformation; MPWTR – Modified Power Transformation; DM – Distribution Mapping].

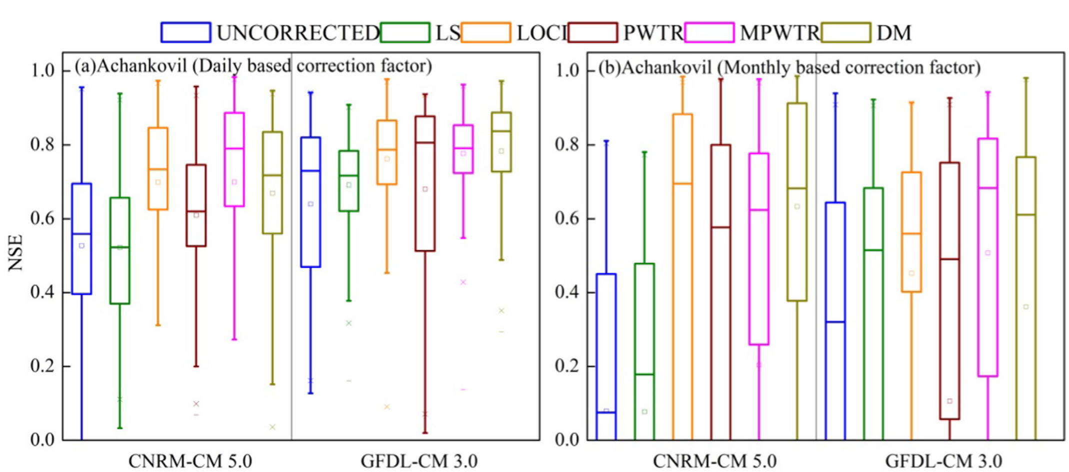

Fig. 6. NSE values in Achankovil basin for (a) daily and (b) monthly correction factors for CNRM-CM5.0 and GFDL-CM3.0 [Uncorrected – Downscaled data; LS – Downscaled data corrected using Linear Scaling; LOCI – Local Intensity Scaling; PWTR – Power Transformation; MPWTR – Modified Power Transformation; DM – Distribution Mapping].

Table 2
Mean absolute error in mm for bias corrected rainfall for CNRM-CM5.0 and GFDL-CM3.0.

Watershed	Model name	Daily					Monthly				
		LS	LOCI	PWTR	MPWTR	DM	LS	LOCI	PWTR	MPWTR	DM
Achankovil	CNRM-CM5.0	84.2	18.8	70.9	27	28.7	110	94.9	80	77.8	57
	GFDL-CM3.0	69.7	31.6	65.6	35	37.7	92.2	113	165	81.1	79
Vaippar	CNRM-CM5.0	18.6	7.6	14.4	7.1	6.5	26.9	9.6	37	13	8.2
	GFDL-CM3.0	29.5	6.8	20.7	15	9.1	9.2	33.1	34	9.1	2.6
Birhidhang	CNRM-CM5.0	129.4	34	143	20	23.7	129	34	143	23.7	20
	GFDL-CM3.0	85.8	28.5	87.9	21	22.3	52	46.6	99	42.6	43
Ken	CNRM-CM5.0	17.6	9.7	34.4	37	11.2	16.9	32	56	27.2	25
	GFDL-CM3.0	42.5	8.9	41.6	8.5	7	16.9	78.5	56	64.9	80
Subarnarekha	CNRM-CM5.0	43.5	18	42.4	26	11.4	20.8	56.3	82	51.2	46
	GFDL-CM3.0	45.3	16.2	46.4	20	7.8	40.7	21	21	14.2	25
Sabarmathi	CNRM-CM5.0	41.9	14	38.7	9.9	9.2	17.4	24.5	35	17.9	10
	GFDL-CM3.0	36.4	9	32.4	5.4	7.8	18.4	37.8	19	4.2	20

correction methods (except LS method), and the climate zone of the basin, the MAE for the methods based on daily correction factor is considerably lower when compared to the MAE for bias adjustment methods based on monthly correction factors. The MAE of the DM method based on daily correction factor is the lowest among all the methods for all the basins. For the GFDL model, the LS method based on monthly correction factor seems to have a lower MAE (better performance) when compared to daily correction factors for all the river basins except Achankovil. However, with the CNRM model, the monthly correction factor of LS method seem to outperform, the daily correction factor of LS method for only three basins. This difference in the performance of LS methods applied with monthly correction factor versus daily correction factor, could be due to the simple linear scaling that was also applied at the same monthly scale for correction (same monthly time scale as the performance statistics). So the mean monthly rainfall with monthly correction factor seemingly performs better than the daily correction factors with rainfall accumulated on a monthly scale. While, this apparent lacuna caused due to the mere adjustment of mean rainfall alone in the case of LS method is overcome to a greater extent by other bias correction methods, especially DM method, which clearly demonstrate the advantages of applying daily bias correction factors over monthly bias correction factors (applied on a daily basis) for better adjustment of RCM simulated rainfall.

5.1.3. Performance statistics of bias correction methods with daily correction factors for India

Statistical evaluation of bias correction methods was carried out for entire India using performance measure NSE estimated for all the 365 days in a year for the entire dataset (various bias correction methods) and uncorrected data with reference to the IMD gridded data. The median NSE value for each grid was plotted across India (Fig. 7) to assess the performance of different bias correction methods. Fig. 7(a) indicates that the uncorrected RCM rainfall deviates significantly from observations over central and north eastern parts of India. All the bias correction methods could improve the RCM simulated rainfall; however, there are significant differences in the level of reducing the bias. The performance of the LS is poor with NSE values below 0.4 for most of the north, northwest and north east portions of India. LS method adjusts only the mean values with respect to IMD data, while the LOCI method which corrects the mean, wet day frequencies and intensities by the application of rainfall threshold prior to application of scaling could correct the RCM data reasonably well except for the some regions in north eastern and western India (Fig. 7(c)). This could be due to the fact that LOCI method does not correct the variance of daily rainfall. It is worthwhile to note from Fig. 7(d) that the performance of PWTR method is poor in the north, north western and south central portions of India which mostly come under semi arid climate zone (except north region) which is subjected to high variability in the wet day frequencies and intensities of rainfall. This is mainly due to the fact that PWTR method which corrects only the mean and variance was not effective in correcting the bias in wet day frequencies and intensities. This lacuna of PWTR method was overcome in this study by carrying out rainfall threshold prior to correction of variance and mean, thereby removing the drizzle effect in the models in modified power transformation method (MPWTR). This greatly helped in improving the performance of corrected rainfall by MPWTR method. Although the MPWTR method obtained an acceptable NSE value greater than 0.6 for most of India, north western and north eastern parts of India which is subjected to very low and very high rainfall respectively has NSE value below 0.4 (Fig. 7(e)). This nonlinear transformation may not perform well for RCM simulations having a larger bias due to more number of highly uncertain extreme rainfall events. The performance of DM method which matches the distribution functions of observations and RCM simulations was good over the entire country (Fig. 7(f)). This method which corrects mean, variance, wet day frequencies and intensities by employing non-linear transformation could obtain an NSE greater than 0.81 across most climatic regions of India.

5.2. Evaluation of hydrological impact of bias correction

The hydrological impact of bias correction of rainfall was evaluated by simulating the streamflow and comparing with the streamflow obtained using observed (IMD) rainfall data. The comparisons were made in terms of median monthly hydrographs, flow duration curves and monthly streamflow quantiles.

5.2.1. Median monthly hydrograph

The median monthly stream flows for the simulations with observed, uncorrected and bias corrected data were estimated. Fig. 8 shows the plots of the median monthly hydrographs for both CMIP5 models in Achankovil, Vaippar and Birhidhang basin. Both the CMIP5 models provided consistent assessments of the median monthly streamflows closer to IMD simulated streamflow values for DM method in Achankovil basin. Although the performance of MPWTR method as indicated by high NSE values (Fig. 8(a1) and (a2)), were comparable to the DM method, the streamflow in Achankovil was simulated better by using DM method for bias correcting the downscaled rainfall data. While, in Vaippar (Fig. 8(b1) and (b2)) all methods except LS and PWTR method, simulated

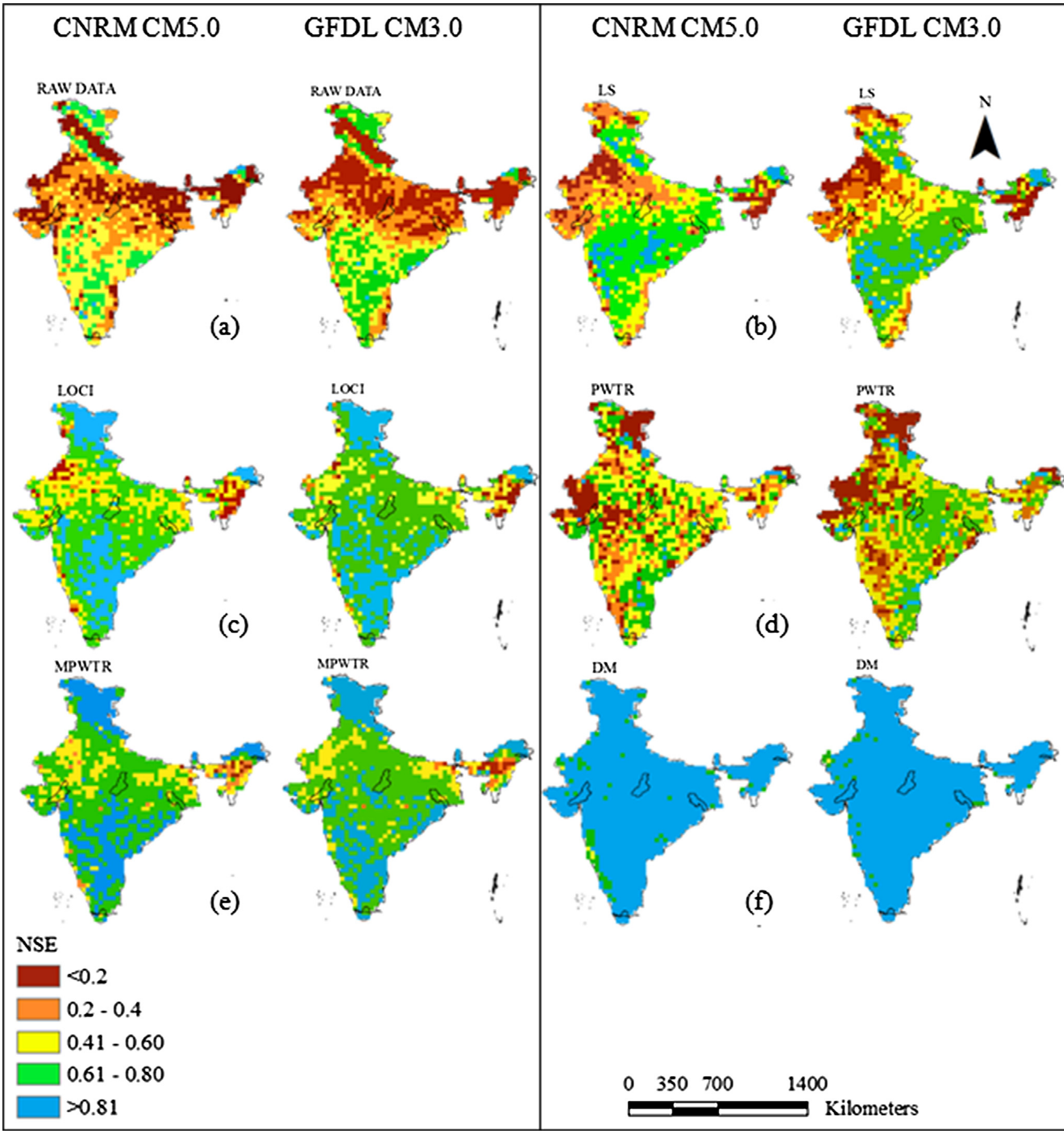

Fig. 7. Median NSE values based on daily correction factors for all India for CNRM-CM5.0 and GFDL-CM3.0 for (a) Uncorrected, (b) LS, (c) LOCI, (d) PWTR, (e) MPWTR, (f) DM [Uncorrected – Downscaled data; LS – Downscaled data corrected using Linear Scaling; LOCI – Local Intensity Scaling; PWTR – Power Transformation; MPWTR – Modified Power Transformation; DM – Distribution Mapping].

the streamflow closer to IMD simulated streamflow values. In Birhidhang basin (Fig. 8(c1) and (c2)), both LS and PWTR methods under estimated the streamflows during the South West monsoon season. LOCI, MPWTR and DM method performed reasonably well in Ken, Subarnarekha, and Sabarmathi basins (Fig. 9). As indicated by the Fig. 9, LS and PWTR method underestimated the streamflow in both Ken and Subarnarekha basin, while in Sabarmathi basin LS and PWTR methods over estimated the streamflows. In catchments located in arid/semi arid climate zones (Sabarmathi and Vaippar), LS and PWTR methods consistently over estimated the streamflow.

It was observed earlier that in the Vaippar basin, there was an overestimation in the number of rainy days with a reasonable estimation of rainfall volume (Fig. 2(b)) and Sabarmathi basin exhibited underestimation of the rainfall volume with a reasonable estimation of rainy days (Fig. 3(f)). This could have resulted in the poor performance of LS and PWTR methods in these two basins. LS and PWTR method which fails to correct the number of rainy days is not suitable for semi-arid catchments with large variability in RCM simulated rainfall. Fang et al. (2015) has also reported that LS method is not ideal for hydrological impact

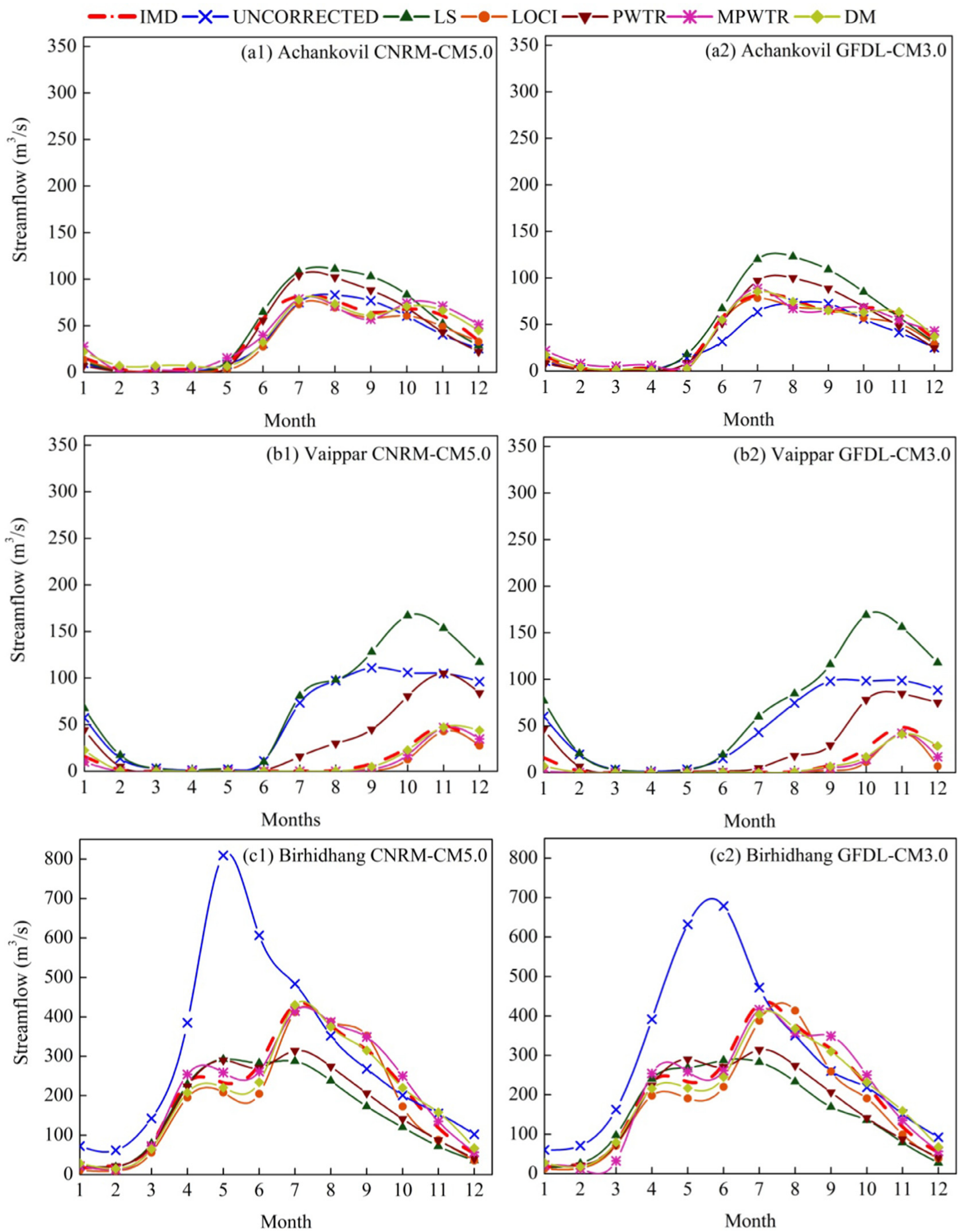

Fig. 8. Median hydrographs of CNRM-CM5.0 and GFDL-CM3.0 for Achankovil (a1 and a2) Vaippar (b1and b2) Birhidhang (c1 and c2) [Uncorrected – Downscaled data; LS – Downscaled data corrected using Linear Scaling; LOCI – Local Intensity Scaling; PWTR – Power Transformation; MPWTR – Modified Power Transformation; DM – Distribution Mapping].

assessments if there is a large variation between observed and RCM simulated rainfall distribution due to its inability in adjusting the statistical properties of rainfall other than its mean. From the Figs. 8 and 9, it could be observed that the median hydrographs estimated with RCM rainfall bias corrected with DM method, compared closely with the median hydrographs simulated with IMD gridded rainfall data. This clearly demonstrates the superior performance of DM method in bias correction. Further, the difference in the magnitude of hydrographs for different bias correction method shows that the choice of bias correction method will have a greater effect on the hydrologic simulations.

5.2.2. Flow duration curve

The effect of bias correction of rainfall data on the simulation of streamflow was further evaluated using monthly flow duration curves to assess the efficacy in simulating different ranges of streamflow and its probability of occurrence. The flow duration curves with monthly average streamflows simulated using IMD, uncorrected and bias corrected rainfall were plotted for all the six catchments (Fig. 10).

From Fig. 10, it can be generally observed that in all the six river basins, the streamflow simulated using rainfall data bias corrected with DM method followed closely the flow

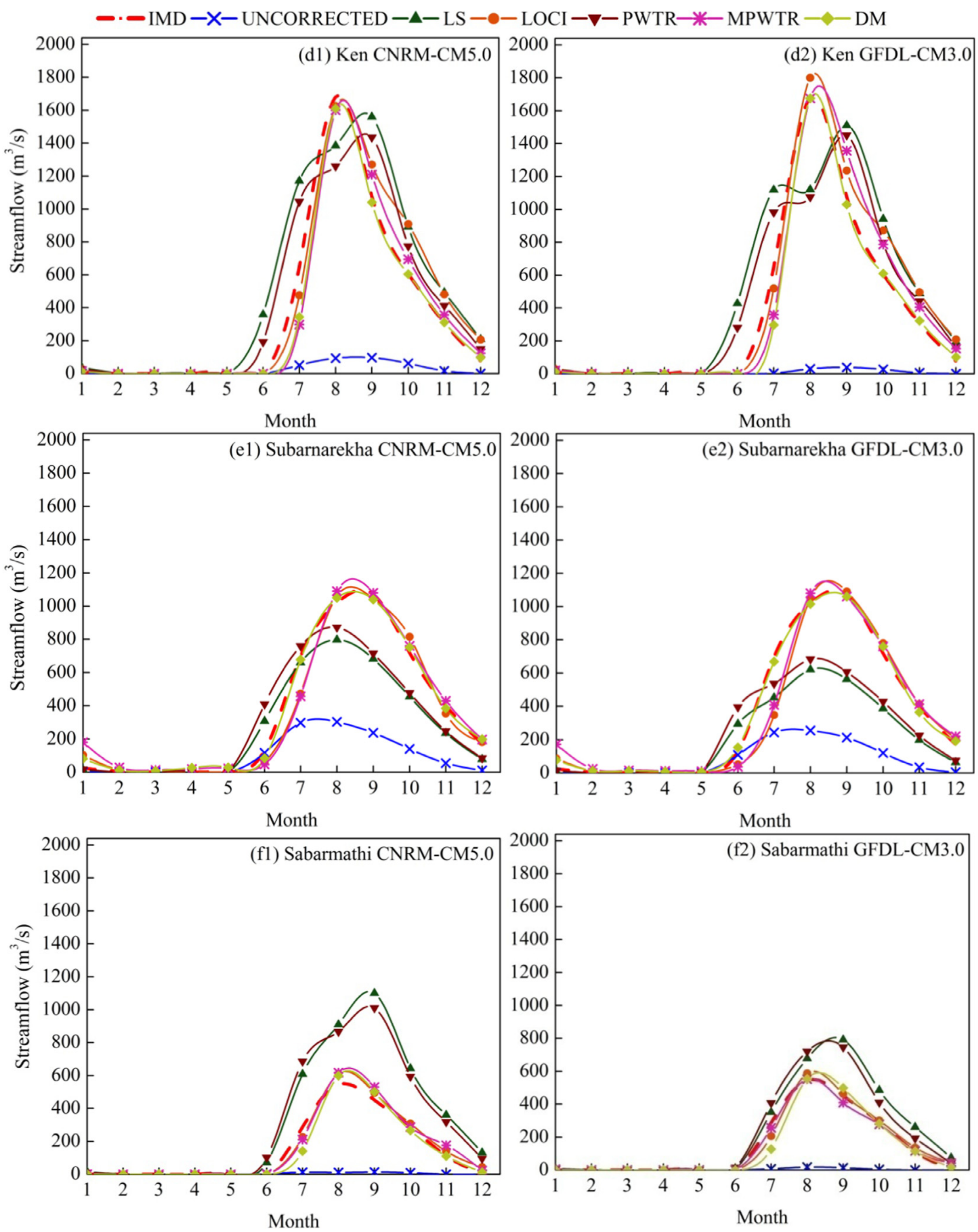

Fig. 9. Median hydrographs of CNRM-CM5.0 and GFDL-CM3.0 for Ken (d1 and d2) Subarnarekha (e1 and e2) Sabarmathi (f1 and f2) [Uncorrected – Downscaled data; LS – Downscaled data corrected using Linear Scaling; LOCI – Local Intensity Scaling; PWTR – Power Transformation; MPWTR – Modified Power Transformation; DM – Distribution Mapping].

distribution simulated using IMD gridded rainfall data. Further, the flow distribution simulated using rainfall data bias corrected with LS and PWTR are not able to replicate the streamflow distribution.

In Achankovil basin which is located in a tropical wet climate, except for the streamflow distribution simulated with LS method, all other methods of rainfall bias correction are able to replicate the streamflow distribution comparable to those simulated with IMD gridded rainfall data (Fig. 10(a1) and (a2)). The overestimation of streamflow by LS method could be due to the overestimation in the number of rainy days by the RCM (Fig. 2(a)), which was not adjusted (wet day frequencies) during the bias correction for this method.

In Vaippar and Sabarmathi basins, which are located in semi-arid and arid climates respectively, the streamflow distribution simulated using LOCI, MPWTR and DM method are able to replicate streamflows very similar to those obtained from IMD gridded data. However, the streamflow distribution with rainfall corrected by LS and PWTR methods have a tendency to overestimate the streamflows. This overestimation of streamflow by LS and PWTR method could be due to the inaccurate simulation in the number of rainy days by the RCM (Fig. 2(b) and (f)) which was not adjusted in terms of wet day frequencies during the bias correction using these methods.

Stream flow distribution simulated with rainfall corrected by LS and PWTR methods have a tendency to underestimate the

streamflow in humid subtropical and tropical climate as indicated for Birhidhang, Ken and Subarnarekha basins (Fig. 10(c1) and (c2); (d1) and (d2); (e1) and (e2)). This underestimation could be due to the difference in wet day frequencies which is not adjusted by the LS and PWTR methods. Hence, the small rainfall quantities is distributed across more number of days resulting in the underestimation of the streamflows. However, the streamflow distribution simulated with rainfall corrected by LOCI, MPWTR and DM method very closely followed the simulation with observed rainfall data, primarily because the wet day frequencies are adjusted in all these three bias correction methods. Due to the adjustment of wet day frequencies, the antecedent moisture in the model is realistically simulated resulting in a better simulation of streamflow distribution comparable to those simulated with observed rainfall data.

5.2.3. Performance statistics for streamflow quantiles

The evaluation of flow duration curve gives an assessment of overall flow distribution within the period of analysis. However, this does not give information about the distribution of streamflow for individual months over the period of analysis. The ability of different rainfall bias correction methods in the simulation of streamflow distribution for individual months in comparison to simulations with IMD gridded data was evaluated using NSE. The performance statistics (NSE) for various quantiles (5, 10 ... 90,

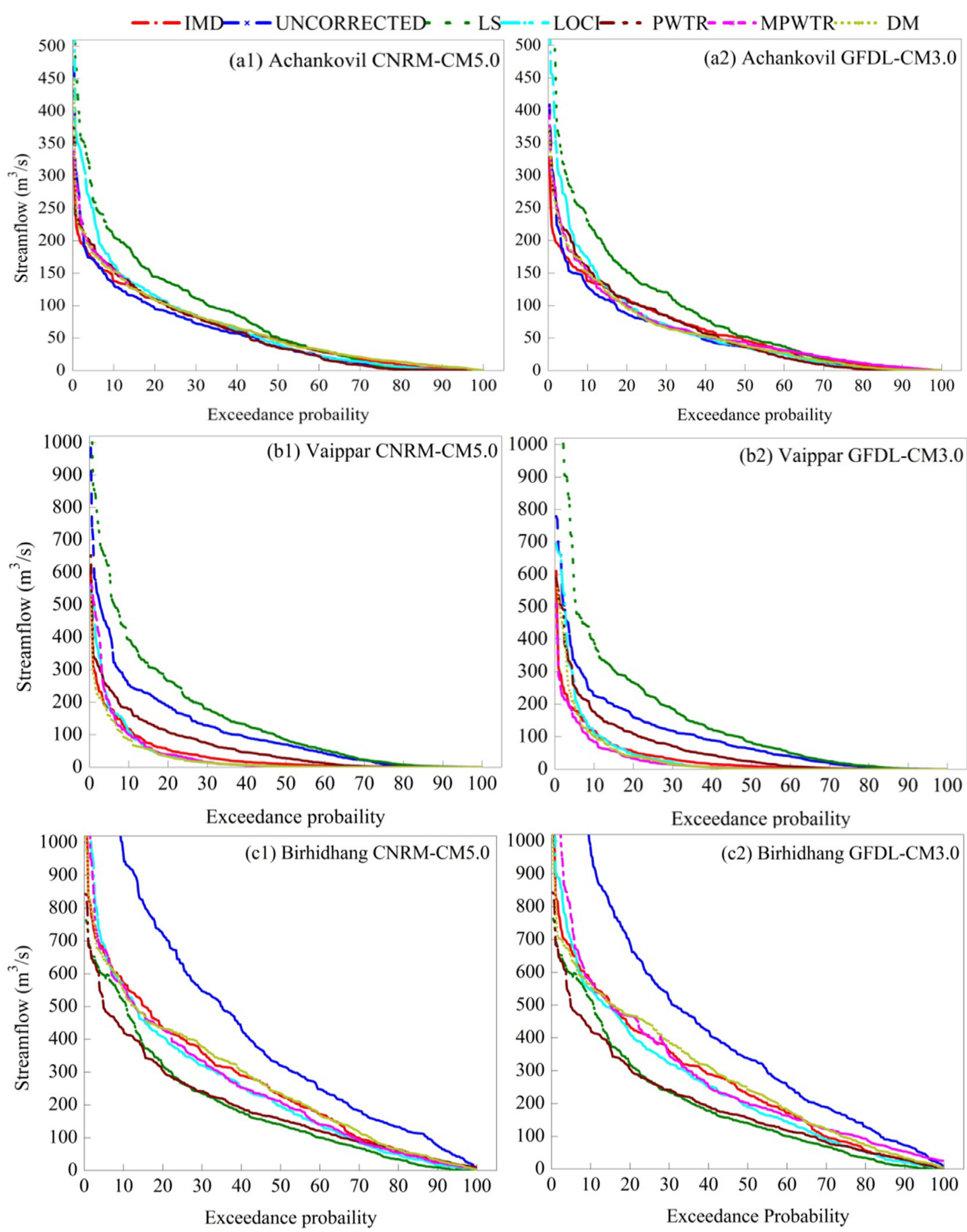

Fig. 10. Flow duration curves of CNRM-CM5.0 and GFDL-CM3.0 for Achankovil basin (a1 and a2) Vaippar (b1 and b2) Birhidhang (c1 and c2) Ken (d1 and d2) Subarnarekha (e1 and e2) and Sabarmathi (f1 and f2) [Uncorrected – Downscaled data; LS – Downscaled data corrected using Linear Scaling; LOCI – Local Intensity Scaling; PWTR – Power Transformation; MPWTR – Modified Power Transformation; DM – Distribution Mapping].

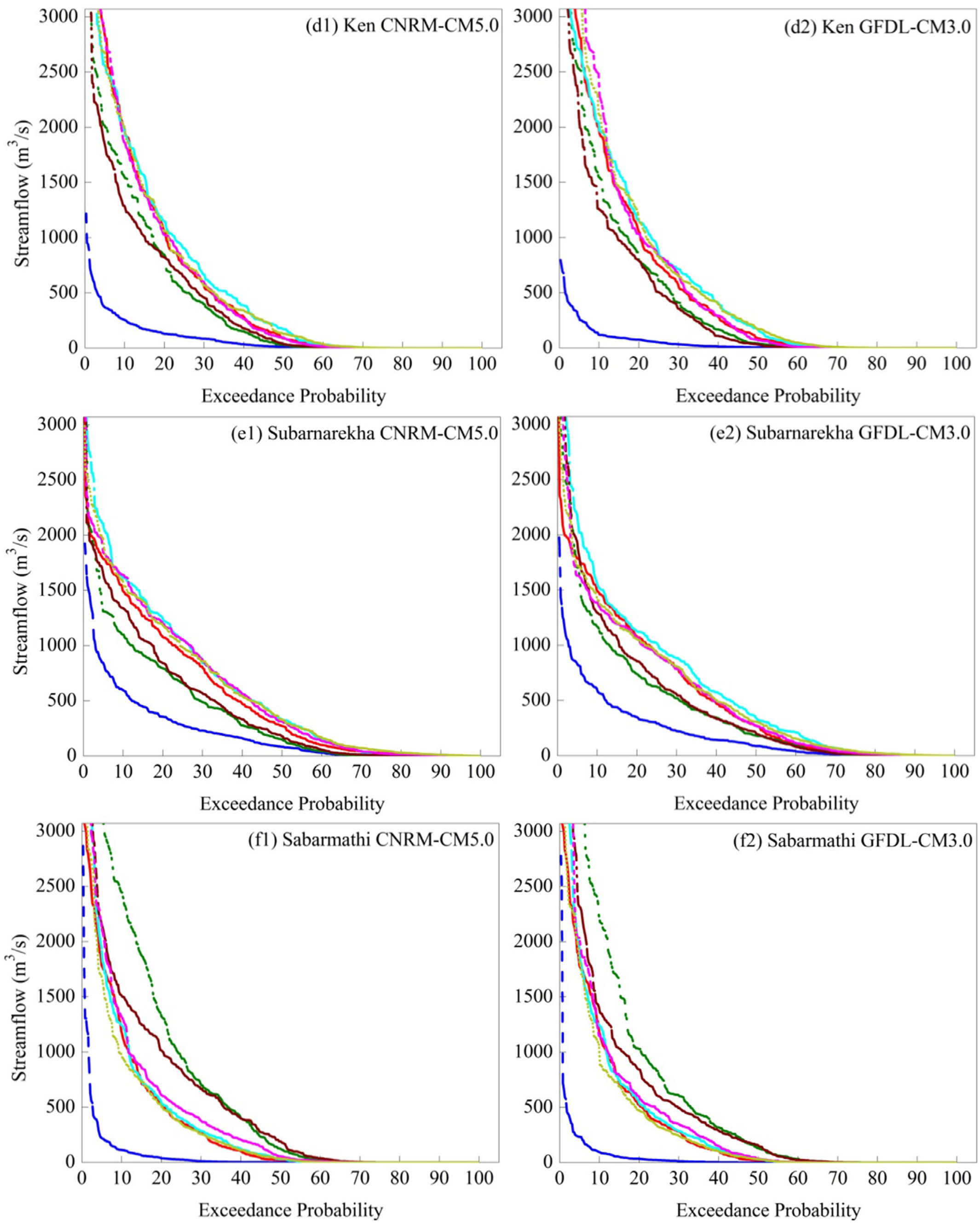

Fig. 10 (*continued*)

95) of streamflow values for each of the twelve months, for uncorrected and bias corrected rainfall compared to streamflow simulated using IMD rainfall values were evaluated for CMIP5 models for all six catchments. Fig. 11 depicts the mosaic plots of NSE values as a measure of different quantiles of streamflow values for each month over the 30 year simulation period.

For Achankovil and Vaippar basin located on the windward and leeward sides of Western Ghats respectively, the performance of the most of the bias correction methods is poor during the months of Jan-Mar (dry season) with high variability in NSE values. These basins experiences winter and pre-monsoon rainfall due to western disturbances driven by westerlies in the southern portion of India (Guhathakurta and Rajeevan, 2008). One plausible interpretation is that due to highly uncertain nature and small amount of rainfall, these disturbances are not adequately simulated by the CMIP5 models (Figs. 3(a) and 2(b)). The high NSE values (NSE > 0.6) during the other months, indicate that the bias correction of rainfall by LOCI and DM methods were able to replicate the distribution of streamflow in each of monsoon months as observed with simulations using IMD gridded data.

In Birhidhang basin located in humid subtropical climate with no dry season, the performance of most of the bias correction methods are reasonably good for majority of the months. This may be due to the reasonable performance of RCM in simulating both the rainy days and rainfall volume over a year (Figs. 2 (c) and 3(c). However, DM method outperformed other methods which is very clear from Fig. 10. In Ken, Subarnarekha and Sabarmathi basins, the performance of LOCI, MPWTR and DM methods

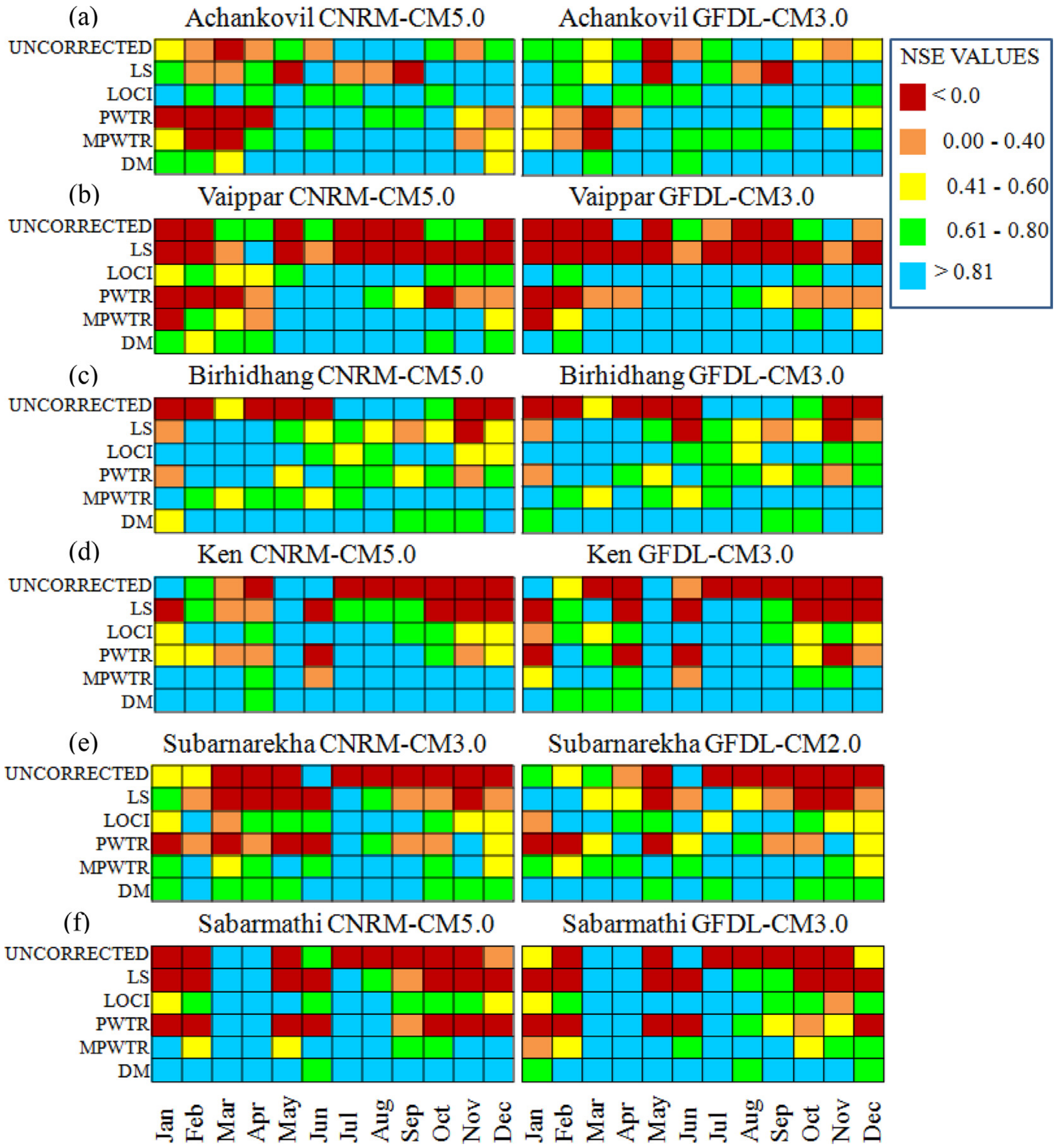

Fig. 11. NSE values for monthly streamflow quantiles for CNRM-CM 5.0 and GFDL-CM 3.0 in (a) Achankovil, (b) Vaippar, (c) Birhidhang, (d) Ken, (e) Subarnarekha and (f) Sabarmathi basin [Uncorrected – Downscaled data; LS – Downscaled data corrected using Linear Scaling; LOCI – Local Intensity Scaling; PWTR – Power Transformation; MPWTR – Modified Power Transformation; DM – Distribution Mapping].

are good compared to other methods with high NSE values for the South West monsoon period, despite underestimation of rainfall by the RCM in these basin as observed earlier (Fig. 3(d)–(f)). This is due to the adjustments in the wet day frequencies and intensities incorporated in these bias correction methods.

In all the basins, the bias correction methods where the wet day frequencies are adjusted viz. LOCI, MPWTR and DM, are able to reasonably replicate the streamflow distribution during the peak monsoon season. However, the performance of LOCI and MPWTR are poor during the dry season. In contrast, even during the dry season, performance of DM method was much superior over other methods in replicating the streamflow distribution of each month. Further, the difference in the NSE values for different bias correction method shows that the choice of bias correction method will have significant effect on the hydrologic simulations.

6. Summary and conclusions

The work presented in this paper contributes towards development and validation of a new methodology for bias correction of RCM simulated rainfall on a daily basis using a sliding window approach. The two best performing CMIP5 models for Indian summer monsoon, CNRM-CM5.0 and GFDL-CM3.0 downscaled using CCAM, are found to have less inter model variability. Although both the models capture very well the Indian summer monsoon (South–West monsoon), there are errors in simulating the peak monsoon month. Both the models simulate the peak monsoon rainfall month one or two months earlier than the observed peaks. Further, in southern parts of India, the retreating monsoon is a major reason for rainfall observed during the months of October to December. However, the rainfall during these retreating monsoon months (rainy days and rainfall magnitude) is not well simulated by the downscaled models. This bias in the rainfall data was corrected using the new methodology proposed in this study.

This work compared the abilities of five rainfall bias correction methods namely linear scaling (LS), local intensity scaling (LOCI), power transformation (PWTR), modified power transformation (MPWTR) and distribution mapping (DM) in correcting RCM simulations for entire India with daily correction factors. The Q-Q plots clearly demonstrated that the bias correction with daily factors

was much superior when compared to methods based on monthly correction factors. The median Nash Sutcliffe efficiency (NSE) based on the Q-Q plots over the 365 days across India depicted, that the DM method was able to bias correct the RCM rainfall data more reliably, irrespective of the seasons or climate zones.

Assessment of streamflow in terms of monthly median hydrographs, flow duration curves and monthly flow distributions clearly demonstrate that the bias correction methods have considerable impact in the simulation of streamflow. The bias correction methods which account for wet day frequency adjustments are able to replicate the streamflow reasonably well (NSE > 0.6). Especially, the LOCI, MPWTR and DM method of bias correction of rainfall was able to replicate the monthly streamflow during the monsoon months. However, the DM method of bias correction of rainfall was able to replicate the monthly streamflow distribution even during the dry months with less number of rainy days. This clearly shows that the DM method based on daily correction factors of RCM rainfall would be a suitable method irrespective of seasons or climate zone for a realistic simulation of hydrologic impacts of climate change.

The main advantage of the proposed new methodology is the ability to preserve the daily characteristics of the rainfall. This method of bias correction based on daily factors enables a smooth transition of daily correction factors, thus, can be considered more appropriate than a sharp transition which occurs in monthly correction factors when applied on a daily time scale. Furthermore, in the proposed sliding window approach, seasonal cycle of the climatology in a region can be replicated well, thereby avoiding climatological discontinuities that occur at the transition between months as in traditional methods. Hence, the proposed method appears to be versatile and can be applied to other regional climatic zones across monsoon influenced countries. It can also be concluded from this study that the DM method is the best daily precipitation bias correction technique. Further, this study also emphasizes the importance of the selection of appropriate bias correction method and applying them on a daily time scale for watersheds to make realistic hydrological assessments.

Although, this study demonstrated that the general distribution patterns of the stream flow are well replicated by the proposed bias correction method, it did not focus on the effectiveness of the bias correction technique in replicating the hydrological extremes. Hence, the future scope of work could assess the effectiveness of the bias correction data in terms of replicating low flow indices and high flow indices. This would help to improve the reliability of climate change impact assessment on different flow regimes that delivers different ecosystem services.

Acknowledgement

We would like to acknowledge CCCR-IITM and IMD, India for providing the data to perform this study.

References

Argüeso, D., Evans, J.P., Fita, L., 2013. Precipitation bias correction of very high resolution regional climate models. Hydrol. Earth Syst. Sci. 17, 4379–4388.

Arnold, J.G., Srinivasan, R., Muttiah, R.S., Williams, J.R., 1998. Large area hydrologic modeling and assessment. Part I: model development. J. Am. Water Resour. Assoc. 34 (1), 73–89.

Ben Alaya, M.A., Chebana, F., Ouarda, T.B., 2015. Probabilistic multisite statistical downscaling for daily precipitation using a Bernoulli-generalized pareto multivariate autoregressive model. J. Clim. 28 (6), 2349–2364.

Bhowmik, R.D., Sankarasubramanian, A., Sinha, T., Patskoski, J., Mahinthakumar, G., Kunkel, K.E., 2017. Multivariate downscaling approach preserving cross-correlations across climate variables for projecting hydrologic fluxes. J. Hydrometeorol. https://doi.org/10.1175/JHM-D-16-0160.1.

Cannon, A.J., 2016. Multivariate bias correction of climate model output: matching marginal distributions and intervariable dependence structure. J. Clim. 29 (19), 7045–7064.

Cannon, A.J., 2017. Multivariate quantile mapping bias correction: an *N*-dimensional probability density function transform for climate model simulations of multiple variables. Clim. Dyn., 1–19

Chen, J., Brissette, F.P., Leconte, R., 2011. Uncertainty of downscaling method in quantifying the impact of climate change on hydrology. J. Hydrol. 401 (3–4), 190–202.

Chen, J., Brissette, F.P., Chaumont, D., Braun, M., 2013. Finding an appropriate bias correction methods in downscaling precipitation for hydrologic impact studies over North America. Water Resour. Res. 49, 4187–4205.

Christensen, J.H., Risnen, J., Iversen, T., 2001. A synthesis of regional climate change simulations a Scandinavian perspective. Geophys. Res. Lett. 28 (6), 1003–1006.

Christensen, J.H., Boberg, F., Christensen, O.B., Lucas-Picher, P., 2008. On the need for bias correction of regional climate change projections of temperature and precipitation. Geophys. Res. Lett. 35, L20709.

Dell' Aquila, A., Calmanti, S., Ruti, P.M., Struglia, M.V., Pisacane, G., Adriana, C., Sannino, G., 2012. Impacts of seasonal cycle fluctuations over the Euro Mediterranean area using a regional earth system model. Climate Res. 52, 135–157.

Dosio, A., Paruolo, P., 2011. Bias correction of the ENSEMBLES high resolution climate change projections for use by impact models: evaluation on the present climate. J. Geophys. Res. 116, D16106.

Eden, J.M., Widmann, M., Grawe, D., Rast, S., 2012. Skill, correction, downscaling of GCM simulated precipitation. J. Clim. 25, 3970–3984.

Everett, J.T., Krovnin, A., Lluch-Belda, D., Okemwa, E., Regier, H.A., Troadec, J.P., Strzepek, K.M., 1996. Fisheries. In: Watson, R.T., Zinyowera, M.C., Moss, R.H. (Eds.), Climate Change 1995: Impacts, Adaptations and Mitigation of Climate Change: Scientific-Technical Analyses – Contribution of Working Group II to the Second Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, pp. 511–537.

Fang, G.H., Yang, J., Chen, Y.N., Zammit, C., 2015. Comparing bias correction methods in downscaling meteorological variables for a hydrologic impact study in an arid area in China. Hydrol. Earth Syst. Sci. 19, 2547–2559.

Fowler, H.J., Blenkinsop, S., Tebaldi, C., 2007. Linking climate change modelling to impacts studies: recent advances in downscaling techniques for hydrological modelling. Int. J. Climatol. 27, 1547–1578.

Gudmundsson, L., Bremnes, J.B., Haugen, J.E., Engen Skaugen, T., 2012. Technical note: downscaling RCM precipitation to the station scale using quantile mapping—a comparison of methods. Hydrol. Earth Syst. Sci. Discuss. 9, 6185–6201.

Guhathakurta, P., Rajeevan, M., 2008. Trends in the rainfall pattern over India. Int. J. Climatol. 28, 1453–1469.

Guyennon, N., Romano, E., Portoghese, I., Salerno, F., Calmanti, S., Petrangeli, A.B., Tartari, G., Copetti, D., 2013. Benefits from using combined dynamical-statistical downscaling approaches – lessons from a case study in the Mediterranean region. Hydrol. Earth Syst. Sci. 17, 705–720.

Hansen, J.W., Challinor, A., Ines, A., Wheeler, T., Moron, V., 2006. Translating climate forecasts into agricultural terms: advances and challenges. Clim. Res. 33, 27–41.

Hasson, S., Lucarini, V., Pascale, S., Böhner, J., 2014. Seasonality of the hydrological cycle in major South and Southeast Asian river basins as simulated by PCMDI/CMIP3 experiments. Earth Syst. Dyn. 5, 67–87.

Haerter, J.O., Hagemann, S., Moseley, C., Piani, C., 2011. Climate model bias correction and the role of timescales. Hydrol. Earth Syst. Sci. 15, 1065–1079.

Hempel, S., Frieler, K., Warszawski, L., Schewe, J., Piontek, F., 2013. A trend-preserving bias correction – the ISI-MIP approach. Earth Syst. Dyn. 4, 219–236.

Hessami, M., Gachon, P., Ouarda, T.B.M.J., St-Hilaire, A., 2008. Automated regression-based statistical Downscaling Tool. Environ. Model. Softw. 23, 813–834.

Hewitson, B.C., Crane, R.G., 1996. Climate downscaling: techniques and application. Clim. Res. 7, 85–95.

Hnilica, J., Hanel, M., Puš, V., 2017. Multisite bias correction of precipitation data from regional climate models. Int. J. Climatol. 37, 2934–2946.

Jeong, D.I., St-Hilaire, A., Ouarda, T.B.M.J., Gachon, P., 2012. Multisite statistical downscaling model for daily precipitation combined by multivariate multiple linear regression and stochastic weather generator. Clim. Change 114, 567–591.

Johnson, F., Sharma, A., 2012. A nesting model for bias correction of variability at multiple time scales in general circulation model precipitation simulations. Water Resour. Res. 48, W01504.

Khalili, M., Nguyen, V.T.V., Gachon, P., 2013. A statistical approach to multi-site multivariate downscaling of daily extreme temperature series. Int. J. Climatol. 33 (1), 15–32.

Kim, K.B., Bray, M., Han, D., 2015. An improved bias correction scheme based on comparative precipitation characteristics. Hydrol. Process. 29, 2258–2266.

King, K., Irwin, S., Sarwar, R., Simonvic, S.P., 2012. The effects of climate change on extreme precipitation events in the Upper Thames River Basin: a comparison of downscaling approaches. Can. Water Resour. J. 37 (3), 253–274.

Lafon, T., Dadson, S., Buys, G., Prudhomme, C., 2012. Bias correction of daily precipitation simulated by a regional climate model: a comparison of methods. Int. J. Climatol. 33, 1367–1381.

Lenderink, G., Buishand, A., Van Deursen, W., 2007. Estimates of future discharges of the river Rhine using two scenario methodologies: direct versus delta approach. Hydrol. Earth Syst. Sci. 11 (3), 1145–1159.

Leander, R., Buishand, T.A., 2007. Resampling of regional climate model output forthe simulation of extreme river flows. J. Hydrol. 332 (3–4), 487–496.

Leander, R., Buishand, T.A., Vandenhurk, B.J.J.M., Dewit, M.J.M., 2008. Estimated changes in flood quantiles of the river Meuse from resampling of regional climate model output. J. Hydrol. 351 (3–4), 331–343.

Li, H., Sheffield, J., Wood, E.F., 2010. Bias correction of monthly precipitation and temperature fields from Intergovernmental Panel on Climate Change AR4 models using equidistant quantile matching. J. Geophys. Res. 115, D10101.
Li, C., Sinha, E., Horton, D.E., Diffenbaugh, N.S., Michalak, A.M., 2014. Joint bias correction of temperature and precipitation in climate model simulations. J. Geophys. Res. Atmos. 119, 13153–13162.
Maurer, E.P., Hidalgo, H.G., Das, T., Dettinger, M.D., Cayan, D.R., 2010. The utility of daily large-scale climate data in the assessment of climate change impacts on daily streamflow in California. Hydrol. Earth Syst. Sci. 14, 1125–1138.
Mehrotra, R., Sharma, A., 2012. An improved standardization procedure to remove systematic low frequency variability biases in GCM simulations. Water Resour. Res. 48, W12601.
Mehrotra, R., Sharma, A., 2015. Correcting for systematic biases in multiple raw GCM variables across a range of timescales. J. Hydrol. 520, 214–223.
Mehrotra, R., Sharma, A., 2016. Multivariate quantile-matching bias correction approach with auto- and cross-dependence across multiple time scales: implications for downscaling. J. Clim. 29, 3519–3539.
Muerth, M.J., St-Denis, B.G., Ricard, S., Velázquez, J.A., Schmid, J., Chaumont, D., Caya, D., Ludwig, R., Turcotte, R., 2013. On the need for bias correction of regional climate scenarios to assess climate change impacts on river runoff. Hydrol. Earth Syst. Sci. 17, 1189–1204.
Murphy, J., 1999. An evaluation of statistical and dynamical techniques for downscaling local climate. J. Clim. 12, 2256–2284.
Nachtergaele, F., Velthuizen, H.V., Verelst, L., 2008. Harmonized World Soil Database. Food Agric. Organ. U. N Technical Report.
Nguyen, H., Mehrotra, R., Sharma, A., 2016. Correcting for systematic biases in GCM simulations in the frequency domain. J. Hydrol. 538, 117–126.
Piani, C., Haerter, J.O., Corpola, E., 2010. Statistical bias correction for daily precipitation in regional climate models over Europe. Theor. Appl. Climatol. 99, 187–192.
Pierce, D.W., Cayan, D.R., Maurer, E.P., Abatzoglou, J.T., Hegewisch, K.C., 2015. Improved bias correction techniques for hydrological simulations of climate change. J. Hydrometeorol. 16, 2421–2442.
Pai, D.S., Latha, S., Rajeevan, M., Sreejith, O.P., Satbhai, N.S., Mukhopadhyay, B., 2014. Development of a new high spatial resolution ($0.25^\circ \times 0.25^\circ$) long period (1901–2010) daily gridded rainfall data set over India and its comparison with existing data sets over the region. MAUSAM 65 (1), 1–18.
PaiMazumder, D., Done, J.M., 2015. The roles of bias correction and resolution in regional climate simulations of summer extremes. Clim. Dyn. 45, 1565–1581.
Salvi, K., Kannan, S., Ghosh, S., 2013. High resolution multi site daily rainfall projections in India with statistical downscaling for climate change impact assessment. J. Geophys. Res. Atmos. 118, 3557–3578.
Schmidli, J., Frei, C., Vidale, P.L., 2006. Downscaling from GCM precipitation: a benchmark for dynamical and statistical downscaling methods. Int. J. Climatol. 26 (5), 679–689.
Sengupta, A., Rajeevan, M., 2013. Uncertainty quantification and reliability analysis of CMIP5 projections for the Indian summer monsoon. Curr. Sci. 105 (12), 1692–1703.
Sennikovs, J., Bethers, U., 2009. Statistical downscaling method of regional climate model results for hydrological modelling. In: 18th World IMACS MODSIM Congress, Cairns, Australia, pp. 3962–3968.
Sharma, D., Das Gupta, A., Babel, M.S., 2007. Spatial disaggregation of bias corrected GCM precipitation for improved hydrologic simulation: ping river basin, Thailand. Hydrol. Earth Syst. Sci. 11 (4), 1373–1390.
Sperber, K.R., Annamalai, H., Kang, I.S., Kitoh, A., Moise, A., Turner, A., Wang, B., Zhou, T., 2012. The Asian summer monsoon: an intercomparison of CMIP5 vs. CMIP3 simulations of the late 20th century. Clim. Dyn. 41, 2711–2744.
Srivastava, A.K., Rajeevan, M., Kshirsagar, S.R., 2009. Development of a high resolution daily gridded temperature data set (1969–2005) for the Indian region. Atmos. Sci. Lett. 10, 249–254.
Sunyer, M.A., Hundercha, Y., Lawrence, D., Madsen, H., Willems, P., Martinkova, M., Vermoon, K., Burger, G., Hanel, M., Kricruciuniene, J., Loukaas, A., Osuch, M., Yucel, I., 2014. Intercomparison of statistical downscaling methods for projection of extreme precipitation in Europe. Hydrol. Earth Syst. Sci. Discuss. 11, 6167–6214.
Themeßl, M.J., Gobiet, A., Leuprecht, A., 2011. Empirical-statistical downscaling and error correction of daily precipitation from regional climate models. Int. J. Climatol. 31, 1530–1544.
Thrasher, B., Maurer, E.P., Mckellar, C., Duffy, P.B., 2012. Technical note: bias correcting climate model simulated daily temperature extremes with quantile mapping. Hydrol. Earth Syst. Sci. 16, 3309–3314.
Teutschbein, C., Seibert, J., 2012. Bias correction of regional climate model simulations for hydrological climate-change impact studies: review and evaluation of different methods. J. Hydrol. 456, 12–29.
Vrac, M., Friederichs, P., 2015. Multivariate—intervariable, spatial, and temporal bias correction. J. Clim. 28, 218–237.
Wilcke, R.A.I., Mendlik, T., Gobiet, A., 2013. Multi-variable error correction of regional climate models. Clim. Change 120, 871–887.
Wood, A.W., Leung, L.R., Sridhar, V., Lettenmaier, D.P., 2004. Hydrologic implications of dynamical and statistical approaches to downscaling climate outputs. Clim. Change 62, 189–216.